

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:22:19 ; Search time 163 Seconds
(without alignments)
1100.962 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLAUSLLAAGVLACGS.....YKFDYDKRIDSTSGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2477	100.0	464	3 AAY79993	Aay79993 Plasmid L
2	1920	77.5	364	2 AAR20108	Aar20108 Protein D
3	1251.5	50.5	358	6 ABB84614	Abb84614 E. coli g
4	1251.5	50.5	359	7 ABO61848	Ab061848 Klebsiell
5	1219.5	49.2	364	7 ADF05118	Adf05118 Bacterial
6	1210.5	48.9	359	6 ABM68758	Abm68758 Phototrab
7	1074	43.4	356	2 AAY52773	Aay52773 Treponema
8	661.5	26.7	451	2 AAY06589	Aay06589 Lipoprote
9	661.5	26.7	451	6 ABU04447	Abu04447 Human exp
10	651	26.3	144	2 AAY02348	Aay02348 Protein D
11	651	26.3	446	2 AAY06590	Aay06590 Lipoprote
12	606	24.5	324	2 AAY02352	Aay02352 A represe
13	606	24.5	324	4 AAG63234	Aag63234 Amino aci
14	606	24.5	324	7 ADB67693	Adb67693 HIV-1 Lip
15	597	24.1	326	2 AAY02354	Aay02354 A represe
16	597	24.1	326	4 AAG63236	Aag63236 Amino aci
17	597	24.1	326	7 ADB67697	Adb67697 HIV-1 Pro
18	594.5	24.0	411	4 AAG63235	Aag63235 Amino aci
19	594.5	24.0	411	7 ADB67695	Adb67695 HIV-1 Lip
20	588.5	23.8	411	2 AAY02353	Aay02353 A represe
21	585.5	23.6	413	4 AAY02355	Aay02355 A represe
22	585.5	23.6	413	2 AAG63237	Aag63237 Amino aci
23	585.5	23.6	413	7 ADB67699	Adb67699 HIV-1 Pro
24	565	22.8	220	2 AAY25375	Aay25375 HPV fusio
25	565	22.8	220	2 AAY25378	Aay25378 HPV fusio

26	565	22.8	220	2 AAY02631	Aay02631 ProtDthr1
27	565	22.8	220	2 AAY02634	Aay02634 Prot.D1/3
28	565	22.8	421	5 AAM50663	Aam50663 Thioresox
29	563	22.7	278	2 AAY25385	Aay25385 HPV fusio
30	563	22.7	278	2 AAY02641	Aay02641 Prot.D1/3
31	561.5	22.7	227	2 AAY25384	Aay25384 HPV fusio
32	561.5	22.7	227	2 AAY02640	Aay02640 Prot.D1/3
33	559.5	22.6	273	2 AAY25376	Aay25376 HPV fusio
34	559.5	22.6	273	2 AAY02632	Aay02632 Prot.D1/3
35	559.5	22.6	371	2 AAY25377	Aay25377 HPV fusio
36	559.5	22.6	371	2 AAY02633	Aay02633 Prot.D1/3
37	558	22.5	220	8 ADL64005	Adl64005 Chimeric
38	558	22.5	220	8 ADL63966	Adl63966 Chimeric
39	558	22.5	220	8 ADL97935	Adl97935 Protein D
40	558	22.5	227	2 AAY25382	Aay25382 HPV fusio
41	558	22.5	227	2 AAY02638	Aay02638 Prot.D1/3
42	558	22.5	383	2 AAY25386	Aay25386 HPV fusio
43	558	22.5	383	2 AAY02642	Aay02642 Prot.D1/3
44	429.5	17.3	797	4 ABG29533	Abg29533 Novel hum
45	381.5	15.4	386	6 ADA34227	Ada34227 Acinetoba

ALIGNMENTS

RESULT 1

AAY79993

ID AAY79993 standard; protein; 464 AA.

AC AAY79993;

DT 15-MAY-2000 (first entry)

DE Plasmid LPD-LB1-III protein sequence.

KW Vaccine; non-typeable Haemophilus influenzae; ntH; infection;
KW chimeric protein; Haemophilus influenzae; P5-like fimbria protein;
KW lipoprotein D; LB1(f); immunogenic; antigenic; otitis media; sinusitis;
KW conjunctivitis; lower respiratory tract infection.

OS Haemophilus influenzae.

OS Synthetic.

XX WO9964067-A2.

XX 16-DEC-1999.

XX 28-MAY-1999; 99WO-US011980.

XX 11-JUN-1998; 98GB-00012613.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX (OHIS) UNIV OHIO STATE RES FOUND.

XX Bakaletz LO, Cohen J, Dequesne G, Lobet Y;

XX WPI; 2000-116457/10.

XX N-PSDB; AAZ91252.

XX Novel antigenic P5-like fimbria subunit peptides used in vaccines against
XX Haemophilus influenzae.
XX Claim 14; Fig 5; 68pp; English.

XX The present invention describes antigenic P5-like fimbria subunit
XX peptides (LB1(f) peptides) of P5-like fimbria proteins from various
XX Haemophilus influenzae strains. The peptides are used for diagnosis,
XX prevention, and treatment of Haemophilus influenzae infections, such as
XX otitis media, sinusitis, conjunctivitis, or lower respiratory tract
XX infection. The peptides may also be used in vaccines against H.
XX influenzae. Antibodies and probes from the present invention can be used
XX for diagnosis of H. influenzae infection. AAY79955 to AAY79993, and
XX AAZ91201 to AAZ91252, represent sequences used in the exemplification of

```

CC the present invention
XX
SQ Sequence 464 AA;

Query Match      100.0%; Score 2477; DB 3; Length 464;
Best Local Similarity 100.0%; Pred. No. 5.1e-207;
Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
   |||||
DB 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60

QY 61 QOADYLEODLAWTKDGRVLVVIHDHFLDGLTDVAKKFPFHRKDGRIYVVDFTLKEIQSLE 120
   |||||
DB 61 QOADYLEODLAWTKDGRVLVVIHDHFLDGLTDVAKKFPFHRKDGRIYVVDFTLKEIQSLE 120

QY 121 MTENFETKDGKQAQVYVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIYPEIKAPW 180
   |||||
DB 121 MTENFETKDGKQAQVYVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIYPEIKAPW 180

QY 161 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
   |||||
DB 161 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240

QY 181 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
   |||||
DB 181 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240

QY 241 AYTDWKETOEDPKGYVWVYNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
   |||||
DB 241 AYTDWKETOEDPKGYVWVYNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300

QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVFTDPDTGVEFL 360
   |||||
DB 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVFTDPDTGVEFL 360

QY 361 KGKSMGCGKAGVALVRSDYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSDYK 420
   |||||
DB 361 KGKSMGCGKAGVALVRSDYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSDYK 420

QY 421 FYEDANGTRDHKKGRHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
   |||||
DB 421 FYEDANGTRDHKKGRHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464

RESULT 2
AAR20108
ID AAR20108 standard; protein; 364 AA.
XX
AC AAR20108;
XX
DT 01-APR-1992 (first entry)
XX
DE Protein D - human IgD-receptor.
XX
KW Immunoglobulin; binding protein; probe; primer; detection; vaccine.
XX
OS Haemophilus influenzae.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Peptide /label= sig_peptide
FT Peptide 16..19
FT Peptide /note= "bacterial lipoprotein consensus sequence"
XX
XX WO9118926-A.
XX
XX 12-DEC-1991.
XX
XX 31-MAY-1990; 90SE-00001949.
XX
XX 31-MAY-1990; 90SE-00001949.
XX
XX (FORS/) FORSGREN A.
XX
XX Forsgren A;
XX

```

```

DR WPI; 1992-007432/01.
XX N-PSDB; AAQ20265.
XX Surface-exposed protein conserved in strains of Haemophilus influenzae -
PT useful in vaccines, and antibodies to protein, are used to detect
PT Haemophilus in samples.
XX
XX Disclosure; Fig 9a-b; 42pp; English.
XX
XX The surface-exposed protein, conserved in many strains of H. influenzae
CC or related Haemophilus species, has a mol. wt. of 42 kD and a binding
CC capacity for human IgD. The DNA sequence can be used to design probes and
CC primers for the detection of H. influenzae and related species. The protein
CC can be used to detect and isolate IgD and to prepare vaccines against
CC H. influenzae
XX
SQ Sequence 364 AA;

Query Match      77.5%; Score 1920; DB 2; Length 364;
Best Local Similarity 99.5%; Pred. No. 1.4e-158;
Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
   |||||
DB 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60

QY 61 QOADYLEODLAWTKDGRVLVVIHDHFLDGLTDVAKKFPFHRKDGRIYVVDFTLKEIQSLE 120
   |||||
DB 61 QOADYLEODLAWTKDGRVLVVIHDHFLDGLTDVAKKFPFHRKDGRIYVVDFTLKEIQSLE 120

QY 121 MTENFETKDGKQAQVYVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIYPEIKAPW 180
   |||||
DB 121 MTENFETKDGKQAQVYVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIYPEIKAPW 180

QY 161 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
   |||||
DB 161 FHHQNGKDIAAETLKVLYKKGDKTDVYVLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240

QY 241 AYTDWKETOEDPKGYVWVYNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
   |||||
DB 241 AYTDWKETOEDPKGYVWVYNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300

QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVFTDPDTGVEFL 360
   |||||
DB 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVFTDPDTGVEFL 360

QY 361 KGIX 364
   |||||
DB 361 KGIX 364

RESULT 3
ABB84614
ID ABB84614 standard; protein; 358 AA.
XX
AC ABB84614;
XX
XX 24-FEB-2003 (first entry)
XX
DE E. coli glycerophosphoryl phosphodiester phosphodiesterase.
XX
XX NARC10; NARC16; cytostatic; immunosuppressive; dermatological; cardiant;
KW antianflammatory; nephrotropic; anti-HIV; nootropic; neuroprotective;
KW antianemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;
KW immunosuppressive; thyromimetic; immunostimulant; antinamic; hypokensative;
KW tranquiliser; neuroleptic; gene therapy; gene mapping; apoptosis;
KW AIDS; cell cycle disruption; programmed cell death regulation;
KW viral infection; nucleosome assembly; phosphate homeostasis;
KW cell cycle regulation; cancer; follicular lymphoma; carcinoma;
KW p53 mutation; graft rejection; hormone-dependent tumour;
KW autoimmune disorder; valvular heart disease;
KW systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis;
KW immune-mediated glomerulonephritis; virus-induced lymphocyte depletion;
XX

```

acquired immunodeficiency syndrome; neurodegenerative disease; stroke; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration; myelodysplastic syndrome; ischaemic injury; myocardial infarction; reperfusion injury; liver disease; idiopathic dilated cardiomyopathy; ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania; myelodysplastic syndrome; central nervous system disorder; anxiety; senile dementia; Huntington's disease; hypertension; schizophrenia; severe bipolar affective disorder; glycerophosphoryl phosphodiester phosphodiesterase.

Escherichia coli.

WO200281516-A2.

17-OCT-2002.

16-JAN-2002; 2002WO-US001098.

16-JAN-2001; 2001US-0262306P.

15-JAN-2002; 2002US-00047855.

(MILL-) MILLENNIUM PHARM INC.

Chiang LW;

WPI; 2003-058503/05.

Novel isolated programmed cell death-related polypeptide, NARC10 and NARC16, useful for treating disorders associated with abnormal apoptotic process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.

Disclosure; Fig 3; 123pp; English.

This invention describes novel cell death-related polypeptides NARC10 and NARC16, located on chromosome 4q11-4q21 and which have cardiant, antiHIV, CC immunosuppressive, dermatological, antiinflammatory, cerebroprotective, CC nootropic, neuroprotective, antianaemic, cardiant, vasotropic, antianemic, CC antididiabetic, immunosuppressive, cytostatic; thyromimetic, nephrotropic, CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and CC neuroleptic activity and can be used in gene therapy. The products of the invention can be used to modulate NARC10 or NARC16 polypeptides or CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene CC regions associated with genetic disease or to associate NARC10 or NARC16 CC with a disease. The polypeptides are also useful for modulating the CC apoptotic process, and are therefore useful for modulating, and treating CC disorders associated with increased apoptosis, inhibition of apoptosis or CC programmed cell death, nucleosome assembly, phosphate homeostasis and the CC cell cycle. Preferably, the products of the invention are useful for CC treating disorders associated with abnormally low rate or abnormally high CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders CC including systemic lupus erythematosus, diabetes, graft rejection, CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral CC infections e.g. infections caused by herpes viruses, virus-induced CC lymphocyte depletion (including acquired immunodeficiency syndrome CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets CC of neurons (including Alzheimer's disease, Parkinson's disease, CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes CC (including aplastic anaemia), ischaemic injuries (including myocardial CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol) CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic CC neutropenia, and myelodysplastic syndromes, central nervous system CC disorders, senile dementia, Huntington's disease, hypertension, CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar CC affective disorder (BP-I). This sequence represents the E. coli CC glycerophosphoryl phosphodiester phosphodiesterase protein described in CC the method of the invention

Sequence 358 AA;

Query Match	50.5%;	Score 1251.5;	DB 6;	Length 358;
Best Local Similarity	64.1%;	Pred. No. 2.7e-100;		
Matches 229;	Conservative	56;	Mismatches 65;	Indels 7; Gaps 2;

Qy	4	KTALSLAAGVLAGCSHSSNMANTQKSDKIIIAHAGSGYLPETHLESKALAFACQA	63
Db	6	KNLSMAIMMSTIVMGSSAMAD-----SNEKIVIAHAGSGYLPETHLPKAWAYACQA	59
Qy	64	DYLEQDLAMTKDGRVVIHDFLDGLTDVAKKPPHRRKDGRIYVIDFTLKEIOSLEMT	123
Db	60	DYLEQDLVMTKDDNLVLHDLVDRVTDVADRRFDRARKDGRYVAIDFTLDEIKSLKFT	119
Qy	124	NFETKDGQAQVYENRPLMKSHRIHTFDEIEIFIOGLEKSTGKVGIIYFEIKAPWPHH	183
Db	120	GFDIENGKKVQTPYGRFPMGKSDPRVHTFEEIEIFVQGLNHSHTGKVGIIYFEIKAPWPHH	179
Qy	184	QNGKIDIAAETLVKLYGKYDKKTDVYLQTFDFPNELKRIKTELLPQMGMDLKLVLQIAYT	243
Db	180	QEGKIDIAAKTLEVLKLYGKYDKKTDVYLQTFDFPNELKRIKTELLPQMGMDLKLVLQIAYT	239
Qy	244	DWKETQSKDPKGYWVNYNYDMFKPGAMAEVYKYADGVGPGWYMLVNRKESKPDNI	303
Db	240	DMNETQKQKPDGSMVNYNYDMFKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG	298
Qy	304	LVKELAQYNVEVHYTVRKDALPEFFTDVNMQYDALLNKSGATGFTDPTGTGVEFL	360
Db	299	MVQDAQKMLVHPYTVRSKPLPEYTPDVNLQYLNKAGVNGLFTDPPDKAVKFL	355

RESULT 4	
ABO61848	
ID	ABO61848 standard; protein; 359 AA.
XX	
AC	ABO61848;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Klebsiella pneumoniae polypeptide seqid 8365..
XX	
KW	Recombinant expression vector; transcription regulatory element;
XX	
KW	Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX	
OS	Klebsiella pneumoniae.
XX	
FN	US6610836-B1.
XX	
PD	26-AUG-2003.
XX	
PF	27-JAN-2000; 2000US-00489039.
XX	
PR	29-JAN-1999; 99US-0117747P.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Breton GL, Osborne M;
XX	
DR	WPI; 2003-895346/82.
XX	
DR	N-PSDB; ACH95399.
XX	
PS	Disclosure; SEQ ID NO 8365; 932pp; English.
XX	
CC	The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This is the amino acid sequence of a Klebsiella pneumoniae polypeptide of the invention
XX	

CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (BPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 451 AA;

Query Match 26.7%; Score 661.5; DB 6; Length 451;
Best Local Similarity 37.5%; Pred. No. 1.1e-48;
Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
QY 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QQADYLEODLANTKQRLVVIHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QQADYLEODLANTKQRLVVIHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFETKDGKQAOVYPRNFRPLWKSHFRHTFDEIFIQGLEKSTGKVKGIYPIKAPW 180
DB 121 MTENFETWLEORSQ-----HCKPEE-----GLE-ARGAELGLV-GAQA 159
QY 181 FHHQNGKQIAETLKVLYKGYDKTDMVYLOTFDFNELKRIKTELLPOMGMDL-KLVQL 239
DB 160 TBEQEAASSSTLVEV-----TLGEVPAESPDPPQSGASSLPTT 201
QY 240 IAYTWKETQE-----KDPKGVYVNYNDMMKPG-----MAEVV-----AD 279
DB 202 MNPVLSQSYEDSSNQEEBGPFPDLESEFOALSRKVAELVHFLLLKYRABEPTVKA 261
QY 280 GVGP---GWTMLVNKEESKPDNIVTPLVKELAQYNVEVHPYTVRKDALPEFTDVNQMY 336
DB 262 MLGSVGVGNWQYFPFVFSKASSLQVFGIELMEVDPIGHLI-----PATCGLSY 313
QY 337 DALLNKSGATGVTFDPDTGVBEFL-KGKSMDDGGKAGVALVRSDYKLY-----NKNSSN 390
DB 314 DGLLGDNQI-----MPKAGLLIIVLAIITAREGDCAPBEKIWEELSVLEFVEGREDSILG 367
QY 391 STLKNLGEHHRARAMDGGKAGVALVRSDYKLYFEDANG-----TRDHKK 433
DB 368 DPKLLTQH-----FVOENYLEYRQVFGSDPACYEFMGLMGRALVETSYKV 413
QY 434 GHRTARTSRDYKFD-----NKRIDSTSGHHHHH 464
DB 414 LHHWVKISGGPHISVPLHEWVLRGEETSGHHHHH 451

RESULT 10
AAY02348
ID AAY02348 standard; protein; 144 AA.
XX
AC AAY02348;
XX
DT 17-OCT-2003 (revised)
DT 09-JUL-1999 (first entry)
DE Protein D, the fusion partner for HIV-1 Nef/Tat fusion proteins.
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

KW vaccine; HIV infection; protein D.
XX
OS Haemophilus influenzae; B.
XX WO9916884-A1.
XX
PD 08-APR-1999.
XX
PF 17-SEP-1998; 98WO-EP006040.
XX
PR 26-SEP-1997; 97GB-00020585.
XX (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Bruck C, Godart SAG, Marchand M;
XX WPI; 1999-302282/25.
DR N-PSDB; AAX35684.
XX
PT HIV Tat or Nef protein linked to a fusion partner.
XX
PS Example 2; Fig 1; 66pp; English.
XX
CC The present sequence represents protein D. This protein acts as the
CC fusion partner for the fusion proteins of th invention, which also
CC comprise HIV-1 Nef or Tat proteins (or derivative). The fusion protein
CC can be used in a vaccine to prevent HIV infection. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 144 AA;
Query Match 26.3%; Score 651; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.7e-48;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QQADYLEODLANTKQRLVVIHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QQADYLEODLANTKQRLVVIHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFET 127
DB 121 MTENFET 127
RESULT 11
AAY06590
ID AAY06590 standard; protein; 446 AA.
XX
AC AAY06590;
XX
DT 26-OCT-1999 (first entry)
XX
DE Lipoprotein D-MAGE-1-His fusion protein.
XX
KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;
KW breast cancer; bladder cancer; lung cancer;
KW head and squamous cell carcinoma; colon cancer; oesophagus carcinoma;
KW vaccine; human.
XX
OS Haemophilus influenzae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
FN WO9940188-A2.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-EP000660.

```

XX 05-FEB-1998; 98GB-00002543.
PR 06-FEB-1998; 98GB-00002650.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Cabezón Silva T, Cohen J, Slaoui WM, Vinals Bassols C;
PI WPI; 1999-494293/41.
DR N-PSDB; AAX87591.
XX
XX New protein derivatives used in cancer vaccine therapy for treating a
PT range of cancers including melanomas, carcinomas and cancers of breast.
XX Example 6; Page 67-68; 72pp; English.
XX
XX The present sequence represents a novel fusion protein composed of
CC lipidated protein D (LPD) of Haemophilus influenzae B, the human MAGE-1
CC tumour-associated antigen and a hexahistidine tail. The invention relates
CC to MAGE proteins fused to an immunological fusion partner such as LPD.
CC The LPD moiety provides the fusion protein with additional exogenous T-
CC cell epitopes and also increase expression levels in E. coli. The lipid
CC tail ensures optimal presentation of the antigen to antigen-presenting
CC cells. The affinity tag facilitates purification. The novel fusion
CC proteins provide vaccines for immunotherapy of melanomas or other MAGE-
CC associated tumours like breast, bladder, lung and non-small cell lung
CC cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
CC carcinoma
XX
XX Query Match 26.3%; Score 651; DB 2; Length 446;
XX Best Local Similarity 100.0%; Pred. No. 8.8e-48; Indels 0; Gaps 0;
XX Matches 127; Conservative 0; Mismatches 0;
XX
XX 1 MDPKTLALSLLAAGVLACSSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
XX 1 MDPKTLALSLLAAGVLACSSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
XX
XX 61 QOADYLEQDLAMTKDGRLLVWIHDPLDGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLE 120
XX 61 QOADYLEQDLAMTKDGRLLVWIHDPLDGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLE 120
XX
XX 121 MTENPET 127
XX 121 MTENPET 127
XX
XX
XX RESULT 12
XX AAY02352
XX ID AAY02352 standard; protein; 324 AA.
XX AC AAY02352;
XX
XX 09-JUL-1999 (first entry)
XX
XX A representative Lipod-Nef-His fusion protein.
XX
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;
XX vaccine; HIV infection; protein D.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9916884-A1.
XX
XX 08-APR-1999.
XX
XX 17-SEP-1998; 98WO-EP006040.
XX
XX 26-SEP-1997; 97GB-00020585.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

```

```

XX Bruck C, Godart SAG, Marchand M;
XX WPI; 1999-302282/25.
DR N-PSDB; AAX35688.
XX
XX HIV Tat or Nef protein linked to a fusion partner.
XX
XX Disclosure; Fig 2; 66pp; English.
XX
XX The present sequence represents a fusion protein comprising Lipod-HIV-1
XX Nef-His. The protein is exemplifies the fusion proteins of the invention.
XX The specification also describes fusion proteins comprising HIV-1 Tat
XX protein. The fusion protein can be used in a vaccine to prevent HIV
XX infection
XX
XX Sequence 324 AA;
XX
XX Query Match 24.5%; Score 606; DB 2; Length 324;
XX Best Local Similarity 37.3%; Pred. No. 4.7e-44;
XX Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;
XX
XX 19 CSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPAQOQADYLEQDLAMTKDGRLL 78
XX 1 CSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPAQOQADYLEQDLAMTKDGRLL 60
XX
XX 79 VTIHDHFDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETKQKQAVYPN 138
XX 61 VTIHDHFDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETKQKQAVYPN 113
XX
XX 139 RFPLKSHFRHTPEDEIEFTQGLEKSTGKKGVIPIKAPWFHONGKDIAETLKLK 198
XX 114 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
XX
XX 199 KYGVDKTDMVYLTQDFNELKRIKTELLPQMGMDLKLVOLIATDKETQKDPKGVWV 258
XX 148 KHGAITSSNTAATN-----AACAWLEAQEEVEVGPV 179
XX
XX 259 NYNYDNWFKCMAEAVVYADGVPGWYMLVNKESKPDNIVTTPLVKELAQYVNEVHPY 318
XX 180 T-----PQVLRPMTYKAAVDLSHFL---KEKGLLEGLHSQR----- 214
XX 319 TVRKDALPEFTDVNQMYDALLNKGATGVFTDF---PDTGVFLKGIKSMGKGAGVA 374
XX 215 --RQDILDLYHTQ-----GYFPDQNYTPGQVRY---PLTFCMCYKLV 255
XX
XX 375 LVRSYDKLYNNKSSNSTLKNLGHHRARADWGKAGVALVRSYDKFYEDANGTRDHKKG 434
XX 256 PVEPD-KVEEANKGENTSLHPVSLH---GMDDPEREVLWRFDRL-----AF 300
XX
XX 435 RHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
XX 301 HHVARELHPEY--FKN---CTSGHHHHH 324
XX
XX RESULT 13
XX AAG63234
XX ID AAG63234 standard; protein; 324 AA.
XX
XX AC AAG63234;
XX
XX 01-OCT-2001 (first entry)
XX
XX Amino acid sequence of a His tagged LipD-Nef of HIV.
XX
XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.
XX
XX Synthetic.
XX Human immunodeficiency virus.
XX
XX Key Location/Qualifiers
XX Peptide 1..109
XX /note= "ProtD fusion partner"

```


Qy 79 VVIHDFLDGLTDVAKKPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 138
 |||||
 Db 61 VVIHDFLDGLTDVAKKPHRRKDGRIYVDFITLKEIQSLEMTENFETMGK----- 113
 Qy 139 RPPLWKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHHQNGKIDAAETLKVLK 198
 |||||
 Db 114 ----W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLE 147
 Qy 199 KGYDCKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEKDPKGYWV 258
 |||||
 Db 148 KHGAITSNTAATN-----AACAWLEAQEEEEVEGPPV 179
 Qy 259 NNYNDMMFKPGAMAEVVKYADGVGPGWMLYNKEESKPDNIYVTPLVKELAQYNVEHPY 318
 |||||
 Db 180 T-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLHSOR----- 214
 Qy 319 TVRKDALPEFTDVNQMYDALLNKGSGATGVFTDF-----PDTGVBFELKGIKSMGDKGAGVA 374
 |||||
 Db 215 ---RQDILDLIWYHTQ-----GYFPDMQNYTPGPGVRY---PLTFGWCYKLV 255
 Qy 375 LVRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKFYEDANGTRDHKKG 434
 |||||
 Db 256 VPEPD-KVEEANKGENTSLHHPVSLH---GMDDPEREVLWRFDSRL-----AF 300
 Qy 435 RHTARTSRDYKFDYDNKRIDSTSGHHHHH 464
 |||||
 Db 301 HVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 15

AA02354
 ID- AA02354 standard; protein; 326 AA.

XX AC AA02354;

XX DT 09-JUL-1999 (first entry)

XX DE A representative Lipod-Nef fusion protein.

XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

XX KW vaccine; HIV infection; protein D.

XX OS Synthetic.

XX OS Human immunodeficiency virus 1.

XX PN WO9916884-A1.

XX PD 08-APR-1999.

XX PF 17-SEP-1998; 98WO-EP006040.

XX PR 26-SEP-1997; 97GB-00020585.

XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX XX WPI; 1999-302282/25.

XX DR N-PSDB; AAX35690.

XX PT HIV Tat or Nef protein linked to a fusion partner.

XX PS Disclosure; Fig 2; 66pp; English.

XX CC The present sequence represents a fusion protein comprising Lipod-HIV-1
 CC Nef. The protein is exemplifies the fusion proteins of the invention. The
 CC specification also describes fusion proteins comprising HIV-1 Tat
 CC protein. The fusion protein can be used in a vaccine to prevent HIV
 CC infection

XX SQ Sequence 326 AA;

Query Match 24.1%; Score 597; DB 2; Length 326;

Best Local Similarity 37.2%; Pred. No. 2.9e-43;
 Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Qy 20 SSSSSNNANTOMKSDKIIIIAHRGASGYLPEHTLESKALAPAAQADYLEQDLAMTKDGRLV 79
 |||||

Db 4 SSSSSNNANTOMKSDKIIIIAHRGASGYLPEHTLESKALAPAAQADYLEQDLAMTKDGRLV 63
 |||||

Qy 80 VVIHDFLDGLTDVAKKPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 139
 |||||

Db 64 VVIHDFLDGLTDVAKKPHRRKDGRIYVDFITLKEIQSLEMTENFETMGK----- 115
 |||||

Qy 140 FPLWKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHHQNGKIDAAETLKVKK 199
 |||||

Db 116 ---W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLEK 150
 |||||

Qy 200 YGYDCKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEKDPKGYWV 259
 |||||

Db 151 HGAIITSNTAATN-----AACAWLEAQEEEEVEGPPVT 182
 |||||

Qy 260 YNYNDMMFKPGAMAEVVKYADGVGPGWMLYNKEESKPDNIYVTPLVKELAQYNVEHPYT 319
 |||||

Db 183 -----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLHSOR----- 216
 |||||

Qy 320 VRKDALPEFTDVNQMYDALLNKGSGATGVFTDF-----PDTGVBFELKGIKSMGDKGAGVAL 375
 |||||

Db 217 -RQDILDLIWYHTQ-----GYFPDMQNYTPGPGVRY---PLTFGWCYKLV 258
 |||||

Qy 376 VRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKFYEDANGTRDHKKG 435
 |||||

Db 259 VPEPD-KVEEANKGENTSLHHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303
 |||||

Qy 436 HTARTSRDYKFDYDNKRIDSTSGHHHHH 464
 |||||

Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

Search completed: June 22, 2005, 11:37:45
 Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:32:20 ; Search time 43 Seconds
(without alignments)
805.515 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLALLAAGVLGCS.....YKFYDNKRIDSTSGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1251.5	50.5	359	4	US-09-489-039A-8365
2	1219.5	49.2	364	4	US-09-543-681A-6003
3	565	22.8	220	3	US-09-485-885-1
4	565	22.8	220	3	US-09-485-885-8
5	563	22.7	278	3	US-09-485-885-21
6	561.5	22.7	227	3	US-09-485-885-19
7	559.5	22.6	273	3	US-09-485-885-4
8	559.5	22.6	371	3	US-09-485-885-6
9	558	22.5	227	3	US-09-485-885-16
10	558	22.5	383	3	US-09-485-885-23
11	381.5	15.4	386	4	US-09-328-352-5514
12	381.5	15.4	503	4	US-09-252-991A-27882
13	269.5	10.9	316	4	US-08-956-171E-5231
14	269.5	10.9	316	4	US-08-781-986A-5231
15	265	10.7	353	3	US-09-134-001C-3115
16	245.5	9.9	383	4	US-09-252-991A-25535
17	201	8.1	260	3	US-09-134-001C-5097
18	182.5	7.4	359	1	US-08-457-997B-2
19	182.5	7.4	359	3	US-08-467-722A-2
20	182.5	7.4	359	4	US-09-451-184-2
21	173	7.0	247	4	US-09-107-532A-3856
22	169.5	6.8	203	4	US-09-710-279-1794
23	149.5	6.0	338	1	US-08-210-394-1
24	148.5	6.0	256	3	US-09-134-001C-4544
25	143.5	5.8	313	4	US-09-107-532A-6446
26	133	5.4	382	4	US-09-134-000C-3829
27	130.5	5.3	256	4	US-09-583-110-5215

ALIGNMENTS

RESULT 1

US-09-489-039A-8365
; Sequence 8365, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8365

; LENGTH: 359

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-8365

Query Match 50.5%; Score 1251.5; DB 4; Length 359;
Best Local Similarity 65.3%; Pred. No. 4.2e-108;
Matches 231; Conservative 48; Mismatches 74; Indels 1; Gaps 1;

Qy	7	ALSLAAGVLACGSSHSNNMANTOMKDKIIIAHRGASGYLPPEHTLESKALAFQAQADYL	66
Db	3	AMKMKLTALMSGMTLSLSSALCFSAATAADKMVIAHRGASGYLPPEHTLPKAKMAYAQGADYL	62
Qy	67	BODLAWTKDGRVVIHDFHLDGLTDVAKKPPHRKDGRIYVIDFTLKEIQSLEMTENPE	126
Db	63	EQDLVMTKDLRLVVDHLDYLDVTDVQAFQARAKDGRFYAIDFTLDEIKSLKFTGEFE	122
Qy	127	TKGQKQAVYVNPRLMKSHFRITPEDIIEFIOGLEKSTCKKGYIYPEIKAPWPHHQNG	186
Db	123	PKNGKNVQTYGCRFPMGKSDPRIHTFEEIEFVQGLNHSHTCKNIGIYPEIKAPWPHHQEG	182
Qy	187	KDIAAETLVKKYGYDKKTDVYVLPDFNFKRIKTELLPQMGMDLKLVIQIAYTQWK	246
Db	183	KDIAAETLVKKYGYTSKQDKVYVLPDFNFKRIKTELLPQMGMDLKLVIQIAYTQWK	242
Qy	247	ETQSKDPKGYVWVNYNDWMPKPGMAEVVYKADGVGPGVYMLVKNKEESKPNIVYVPLVK	306
Db	243	ETQSKDPKGYVWVNYNDWMPKPGMAEVVYKADGVGPGVYMLVKNKEESKPNIVYVPLVK	301
Qy	307	ELAQYVVEHYTYVRKDALPEFTFDVNMQYDALLNKSATGVTDFDPTDGYEFL	360
Db	302	EAHSHLQVHYTYVRADQLPEYATNVNQLYDVLNQAQVGDGLFTDFDPAQVQL	355

RESULT 2

```
US-09-543-681A-6003
; Sequence 6003, Application US/09543681A
; Patent No. 6505709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6003
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6003

Query Match 49.2%; Score 1219.5; DB 4; Length 364;
Best Local Similarity 64.7%; Pred. No. 4.1e-105; Indels 5; Gaps 3;
Matches 229; Conservative 50; Mismatches 70;

Qy 8 LSLAAGYLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLE 67
Db 12 LKPLVAGVLLTSL--LSIIA--QAASDKVIAHRGASGYLPEHTLPKALAYAQADYLE 67
Qy 68 QDLAMTKDGLRVVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPET 127
Db 68 QDLVMTKDNQLVHLHDYLDRTVDAERYPVRRKDGRIYVDFTLDEIKLKEPTGFDI 127
Qy- 128 KDGQAQVYVNRFPPLKSHRHTFEDIEFIQLEKSTGKKGVIYPIEIKAPWFHHQNGK 187
Db 128 VDGKKVQSYVNRFPFGKSDFRHTFQEEIEFIQGLNKSTGQDGIYPIEIKAPWFHEQEGK 187
Qy 188 DIAAETLKVLYGDKYKTDVYLTQDFDNELKRIKTELLPQMGNDLKLVLIIATYTDWKE 247
Db 188 DITKKVLEVLQXGYTQKSDNVYLSQDFPDNLKRIKTELLPQMGNDLKLVLIIATYTDWNE 247
Qy 248 TOEKDPKGYVNNYNDWMPKFGAMAEVVKYADGVPQGMVLYNKEKPDNIVTPLVKE 307
Db 248 TYEQPDGTWNTSYNDWMPKFGAMKEIATYADGIGPDYHMLV-BEDSTDPDKITLTGMAAD 306
Qy 308 LAQYVNEVHPVTVRKDALPEFTDVMYDALLNKGATGVFTDPDGTGVFLK 361
Db 307 ARANKLTIHPFTVRVDKLPKYAKOGDQLYDIYNQAGAGVFTDPDGLGVKFLQ 360

RESULT 3
US-09-485-885-1
; Sequence 1, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-1

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 44; Gaps 2;
Matches 117; Conservative 3; Mismatches 8;

Qy 20 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLRV 79
Db 4 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLRV 63
Qy 80 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETK----- 128
Db 64 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETMAMHGDTPTLHE 123
Qy 129 -----DGKQAQVYVNRFPPLKSHRHTF 152
Db 124 YMLDLQPETTDLYCYEQQLNDSSEDEIDGPAGQAEPR-----AHYNIIVTF 170

RESULT 4
US-09-485-885-8
; Sequence 8, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-8

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 44; Gaps 2;
Matches 117; Conservative 3; Mismatches 8;

Qy 20 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLRV 79
Db 4 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLRV 63
Qy 80 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETK----- 128
Db 64 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETMAMHGDTPTLHE 123
Qy 129 -----DGKQAQVYVNRFPPLKSHRHTF 152
Db 124 YMLDLQPETTDLYCYEQQLNDSSEDEIDGPAGQAEPR-----AHYNIIVTF 170

RESULT 5
US-09-485-885-21
; Sequence 21, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
```



```
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-6

Query Match      22.6%; Score 559.5; DB 3; Length 371;
Best Local Similarity 91.1%; Pred. No. 1.2e-43;
Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

Qy 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFETKDGKQ-AQVYPN 138
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFETMAMFQDQERPR 123

Qy 139 RFP 141
Db 124 KLP 126
```

```
RESULT 9
US-09-485-885-16
; Sequence 16, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezón Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-16

Query Match      22.5%; Score 558; DB 3; Length 227;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

Qy 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 127
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 111
```

```
Query Match      22.5%; Score 558; DB 3; Length 227;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

Qy 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 127
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 111
```

```
RESULT 10
US-09-485-885-23
; Sequence 23, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezón Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-23

Query Match      22.5%; Score 558; DB 3; Length 383;
Best Local Similarity 100.0%; Pred. No. 1.8e-43;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

Qy 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 127
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 111
```

```
RESULT 11
US-09-328-352-5514
; Sequence 5514, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5514
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5514

Query Match      15.4%; Score 381.5; DB 4; Length 386;
Best Local Similarity 31.6%; Pred. No. 5e-27;
Matches 125; Conservative 51; Mismatches 145; Indels 75; Gaps 16;
```

```
Qy 6 LALSLAAGVLGAGCSSH---SSNMANTQMSDKI-IIAHKGASGYLPEHTLESKALAFQA 61
Db 13 LCLSLI8---LVGNCDDDKTETTTPTPEYQPKLLVVGHRGASALREHTLASQKAI 69

Qy 62 QADYLEQDLAMTKDGLVVIHDFLDGLTDVAKKFPHRHRK-----DG----RYVIDFT 112
Db 70 GADFIEDLVSTKDGVLVARHENEIGGTTNVSTLSQFADRKTKNIDGVDLTGWFTEDFT 129

Qy 113 LKEIOSLEMT-----NFKTKGKQAOVYPNRPPLWKSHFRHTTFDEFEFTIOGLEKSTCKK 169
Db 130 LSELQQLKARERIPFPRPANTYNDLYP-----VPTLEQIIELAEANYKTKGI 178
```

QY 170 VGIYPEIKAP-WFHONGKDIAAE--TLKVLKKGVDKKTDMVYLQTFDFNELKRIKTEL 226
Db 179 IGLYIETKHTPTFKQN---LAMEDTLTKLAKYKTRDIAVPYILQSFEVQNLKDKREL 235
QY 227 LPQMGMDLKLVLQIAYTDKQKDPKGYWYNNYDMFKPGMAEVVKKYADGVGP--G 284
Db 236 --DLHKTLKHAQIIQLYDSKTSRPADVFESGDTKYADLATQAQGLKDVAKYANGVGP 293
QY 285 WTMLVNKESKPDNIYVYPLVKELAQNVVHPYTVRKD----- 323
Db 294 YILTFNNDGSYKT---STFISDAHTAGLKVHPYFRPENNFPLAPLKCSPOKPAERCPT 349
QY 324 -ALPEFFTDVNMQDALLNKCATGVTFDPTDGV 358
Db 350 GALKEF-----EAY-----FKAGVDGVFTDDPALGRE 376

RESULT 12

US-09-252-991A-27882
; Sequence 27882, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27882
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27882

Query Match 15.4%; Score 381; DB 4; Length 503;
Best Local Similarity 30.5%; Pred. No. 8.4e-27;
Matches 119; Conservative 64; Mismatches 143; Indels 64; Gaps 13;

QY 7 ALSLLAAGVLACGSHSNM-----ANTQM-----KSKIIIAHAGSGLPEHTL 52
Db 133 ACLMLAAGLASAAHADTGNVQRAIDWAQQAQSVRHPAGHERSPLVIAHAGSGLPEHTL 192
QY 53 ESKALAFQAQADYLEQDLAMTKDGLVVIHDFLDGLTDVAK--KFPHRHK---DG--- 104
Db 193 GAYALLAVMAGDYVEPDLVMTDRGKLVARHNDNELGLTTDVAQHPEFADRKTKQKVDGVEL 252
QY 105 -RYVVYIDFTLKEIOSLEMTENFTKDGQAQVYVNPFRPLWKSHPRIHTFEDEIFIOGLE 163
Db 253 TGWFSDEFTLAELKTLRAIERIPT-----IRPGNRL-DGTFTPIQLQEIIDVKSQ 304
QY 164 KSTGKKGVIYPIKAPWFHONGKDIAAETLKLKKGVDKKTDMVYLQTFDFNELKRIK 223
Db 305 ISQORTIGLYPEIKHGTFRQLGAMERPLVNTLHRNGYLGPRAPVFIIQSFEVNNLKEK 364
QY 224 TELLPMGMMDLKLVLQIAYTDKQKDPKGYWYNNYDMFKPGMAEVVKKYADGVGP 283
Db 365 RL-----TGIRLVQL-----YGSQPYDQQAAGGSLTYAEMATAGLRQVARYAYGVP 413
QY 284 GMYLVNKEES---KPDNIYVYPLVKELAQNVVHPYTVR-----KDALPE 327
Db 414 DKSIVIPRDANGNLGP-----TRFVGDAAHAGLKVHPYTFRAENSLPAEFSRSDGNPQ 468
QY 328 FTFDVMQYDALLNKSGATGVTFDPTDGV 357
Db 469 SRGDLAGEIRAVLD-AGIDGLFSDQPDVAV 497

RESULT 13

US-08-956-171E-5231
; Sequence 5231, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5231:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5231:
US-08-956-171E-5231

Query Match 10.9%; Score 269.5; DB 4; Length 316;
Best Local Similarity 27.5%; Pred. No. 1e-16;
Matches 103; Conservative 48; Mismatches 116; Indels 107; Gaps 16;
QY 8 LSLAAGVLACGSHSNM-----TQMKSCKI-IIAHAGSGLPEHTLESKALA 58
Db 30 LSVPTAG-----AEQTQIANKPOAIQWHTNLNTERFTTIAHAGSGLPEHTFOAYDKS 84
QY 59 FAQ-QADYLEODLAWTKDGLVVIHDFLDGLTDVAKKFPHRHKDGRVYVYIDFTLKEIQ 117
Db 85 HNELKASYIEDLQRTKDGHLVAMHDETQVNTT-----NNGHKVEDYTLDELK 132
QY 118 SLEMTENFTKDGQAQV-YENRFPPLWKSHPRIHTFEDEIFIOGLEKSTGKKGVIYPI 176
Db 133 QLDAGSFWNKYPKYARASYKNA-----KVPTLDEILE-----RYGNANYIET 177
QY 177 KAPFHHQNGKDIAAETLKLKKG-----GYDKKTDVMVYLQTFDFNELKRIKTELLPQMG 232
Db 178 KSPDVY-----PGMEEQLLASLKKHLLNNKLNKGNHVMIQSFDSLSKKIHRQ----- 226
QY 233 DLKLVQLIAYTDKQKDPKGYWYNNYDMFKPGMAEVVKKYADGVGPQWYMLVNKE 292
Db 227 -NKHVPLVKVDRKGELOQ-----FNDQRLKRSYAIIGLGP----- 262

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:40:51 ; Search time 159 Seconds
(without alignments)
1120.544 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLALLAAGVLGCS.....YKFYDKRIDSTSGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1251.5	50.5	358	14 US-10-047-855-12	Sequence 12, Appl
2	661.5	26.7	451	16 US-10-473-127-1113	Sequence 1113, Ap
3	651	26.3	144	17 US-10-687-060-7	Sequence 7, Appl
4	606	24.5	324	14 US-10-203-013-15	Sequence 15, Appl
5	606	24.5	324	17 US-10-687-060-15	Sequence 15, Appl
6	606	24.5	324	17 US-10-485-048-15	Sequence 15, Appl
7	597	24.1	326	14 US-10-203-013-19	Sequence 19, Appl
8	597	24.1	326	17 US-10-687-060-19	Sequence 19, Appl
9	597	24.1	326	17 US-10-485-048-19	Sequence 19, Appl
10	594.5	24.0	411	14 US-10-203-013-17	Sequence 17, Appl
11	594.5	24.0	411	17 US-10-687-060-17	Sequence 17, Appl

12	594.5	24.0	411	17	US-10-485-048-17	Sequence 17, Appl
13	585.5	23.6	413	14	US-10-203-013-21	Sequence 21, Appl
14	585.5	23.6	413	17	US-10-687-060-21	Sequence 21, Appl
15	585.5	23.6	413	17	US-10-485-048-21	Sequence 21, Appl
16	565	22.8	420	13	US-10-000-903-1	Sequence 1, Appl
17	565	22.8	420	13	US-10-000-903-8	Sequence 8, Appl
18	565	22.8	420	17	US-10-899-771-1	Sequence 1, Appl
19	565	22.8	420	17	US-10-899-771-8	Sequence 8, Appl
20	565	22.8	421	14	US-10-296-770-7	Sequence 7, Appl
21	563	22.7	428	13	US-10-000-903-21	Sequence 21, Appl
22	563	22.7	428	17	US-10-899-771-21	Sequence 21, Appl
23	561.5	22.7	427	13	US-10-000-903-19	Sequence 19, Appl
24	561.5	22.7	427	17	US-10-899-771-19	Sequence 19, Appl
25	559.5	22.6	427	13	US-10-000-903-4	Sequence 4, Appl
26	559.5	22.6	427	17	US-10-899-771-4	Sequence 4, Appl
27	559.5	22.6	427	13	US-10-000-903-6	Sequence 6, Appl
28	559.5	22.6	427	17	US-10-899-771-6	Sequence 6, Appl
29	558	22.5	427	13	US-10-000-903-16	Sequence 16, Appl
30	558	22.5	427	17	US-10-899-771-16	Sequence 16, Appl
31	558	22.5	427	13	US-10-000-903-23	Sequence 23, Appl
32	558	22.5	427	17	US-10-899-771-23	Sequence 23, Appl
33	360.5	14.6	424	15	US-10-424-599-282063	Sequence 282063,
34	360	14.5	393	14	US-10-156-761-14312	Sequence 14312, A
35	347.5	14.0	391	16	US-10-425-115-195137	Sequence 195137, A
36	343.5	13.9	391	16	US-10-767-701-44879	Sequence 44879, A
37	326.5	13.2	314	15	US-10-282-122A-46840	Sequence 352984,
38	321	13.0	416	16	US-10-425-115-352984	Sequence 45439, A
39	308	12.4	313	15	US-10-282-122A-45439	Sequence 69756, A
40	296	11.9	385	15	US-10-282-122A-69756	Sequence 16, Appl
41	271	10.9	54	14	US-10-152-031-16	Sequence 70281, A
42	269.5	10.9	309	15	US-10-282-122A-70281	Sequence 5231, Ap
43	269.5	10.9	316	8	US-08-781-986A-5231	Sequence 5231, Ap
44	269.5	10.9	316	15	US-10-329-624-5231	Sequence 70498, A
45	265	10.7	349	15	US-10-282-122A-70498	

ALIGNMENTS

RESULT 1
US-10-047-855-12
; Sequence 12, Application US/10047855
; Publication No. US20030165863A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Lillian Wei-Ming
; TITLE OF INVENTION: NARC10 and NARC16, Programmed Cell
; TITLE OF INVENTION: Death-Associated Molecules and Uses Thereof
; FILE REFERENCE: 35800/242056
; CURRENT APPLICATION NUMBER: US/10/047.855
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,306
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli K12
US-10-047-855-12

Query Match 50.5%; Score 1251.5; DB 14; Length 358;
Best Local Similarity 64.1%; Pred. No. 1.7e-93;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

Qy	4	KTALSLAAGVLGCS	SHSNMANTQMSKIIIAHRGASGYLPEHTLESKALAFQA	63
Db	6	KNLSMAIMSTVINGSSAMAAD	-----SNEKVIARHGASGYLPEHTLPKAMAYAQGA	59
Qy	64	DYLBQDLAMTKDGRV	LWTHDHLDTVAKKFPHRUKDGRVYVDFTLKBIQSLWTE	123
Db	60	DYLBQDLVMTKDDNL	VLDHLDVTDVDFDRKDKGRYVDAIDFTLDEIKSLKETE	119
Qy	124	NFETKDGKQAQVY	PNRPFMLWKSHPRIHTFEDEIEFIQGLEKSTGKKVGIPEIKAPWFHH	183

Db 120 GFDIENGKKVQTPGPFPMGKSDPRVHTFEEBIEFVQGLNHTGKNIGIYEIKAPWFH 179
Qy 184 QNGKDIAAETLVKLYGYDKKTDWVYLOTDFDNELKRIKTELLPOMGMDKLVLQIAYT 243
Db 180 QEGKDIAAETLVKLYGYDKKTDWVYLOTDFDNELKRIKTELLPOMGMDKLVLQIAYT 239
Qy 244 DWKETQEKDPKGYWVYNNYDMFKGMAEYVVKVADGVGPGWYMLVKNKEESKPDNIIVTP 303
Db 240 DWNETQQQKPGSGWVYNNYDMFKGMAEYVVKVADGVGPGWYMLVKNKEESKPDNIIVTP 298
Qy 304 LVKELAQNVVEHPYTVRKDALPEPFTDVNQMYDALLNKGATGVFTDPDPTGVVEFL 360
Db 299 MVQDAQKNKLVHPYTVRKDALPEPFTDVNQMYDALLNKGATGVFTDPDPTGVVEFL 355

RESULT 2

US-10-473-127-1113
; Sequence 1113, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1113
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1113

Query Match 26.7%; Score 661.5; DB 16; Length 451;
Best Local Similarity 37.5%; Pred. No. 3.3e-45;
Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
Qy 1 MDPKTLALSLLAAGVLACGSSHSSNMANTQMSKIIIAHKGAGSYLPEHTLESKALAPA 60
Db 1 MDPKTLALSLLAAGVLACGSSHSSNMANTQMSKIIIAHKGAGSYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLVHDFDLGLTDVAKKFPHRHKKDGRYVYVDFTLKEIOSLE 120
Db 61 QOADYLEODLAWTKDGRVLVHDFDLGLTDVAKKFPHRHKKDGRYVYVDFTLKEIOSLE 120
Qy 121 MTENFETDLEQRSQ-----HCKPPE-----GLE-ARGAELGLV-GAQAAPA 159
Db 121 MTENFETDLEQRSQ-----HCKPPE-----GLE-ARGAELGLV-GAQAAPA 159
Qy 181 FHHQNGKDIAAETLVKLYGYDKKTDWVYLOTDFDNELKRIKTELLPOMGMDL-KLVQL 239
Db 160 TEEQEAASSSTLVEV-----TLGEVPAAESPDPPQSGASSLPTT 201
Qy 240 IAYTDWKEQTQ-----KDPKGYWVYNNYDMFKPG-----AAAEVW-----AD 279
Db 202 MNYPLWSQSYEDSSNQEEGSGTSPDLESEFQAALSRKVAELVHLLKYPAREPVTYKAE 261
Qy 280 GVGP---GWYMLVKNKEESKPDNIIVTPLVKSLAQNVVEHPYTVRKDALPEFTDVNQMY 336
Db 262 MLGSVVGNWQVFFPFIKSSSSQLVFGIELMEVDPIGHLIYI-----FATCLGLSY 313

Qy 337 DALLNKSGATGVFTDPDPTGVVEFL-KGIKSMGGKAGVALVRSYDKLY-----NKNSSN 390
Db 314 DGLGDNQI-----MPKAGLLIIVLAIAREGDCAPEEKIWEELSLEVEFEGREDSILG 367
Qy 391 STLNKLGHEHRARAMDGGKAGVALVRSYDKFYEDANG-----TRDHKK 433
Db 368 DPKKLLTQH-----PVQENYLEYRQVPGSDPACYBFLMGPRALVETSYVKV 413
Qy 434 GRHTARTSRSDYKFDYD-----NKRIDSTSGHHHHH 464
Db 414 LHHMVKISGPHISYPLPHEWVLRGEGETSGHHHHH 451

RESULT 3

US-10-687-060-7
; Sequence 7, Application US/10687060
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/BP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pichia pastoris
US-10-687-060-7

Query Match 26.3%; Score 651; DB 17; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.6e-45;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPKTLALSLLAAGVLACGSSHSSNMANTQMSKIIIAHKGAGSYLPEHTLESKALAPA 60
Db 1 MDPKTLALSLLAAGVLACGSSHSSNMANTQMSKIIIAHKGAGSYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLVHDFDLGLTDVAKKFPHRHKKDGRYVYVDFTLKEIOSLE 120
Db 61 QOADYLEODLAWTKDGRVLVHDFDLGLTDVAKKFPHRHKKDGRYVYVDFTLKEIOSLE 120
Qy 121 MTENFET 127
Db 121 MTENFET 127

RESULT 4

US-10-203-013-15
; Sequence 15, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1e1 Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5

```
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-15

Query Match      24.5%; Score 606; DB 14; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60

QY 79 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYVPN 138
DB 61 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYVPN 113

QY 139 RPLWKSHFRHTFEDEIEFIOGLEKSTGKVGIYPEIKAPWFHONGKIDIAAETLKVLL 198
DB 114 -----W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLE 147

QY 199 KYGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAIYTDWKETQKPKGYWV 258
DB 148 KHGAITSNTAATN-----AACAWLEAQEEBEVGPV 179

QY 259 NNYDWMFKPGMAEVKYGADVGPGWYMLVNKESKPDNIYVPLVKELAQYNVVHPY 318
DB 180 T-----PQVPLRPMTYKAADLSHPL---KEKGGLEGLHSQR----- 214

QY 319 TVRKDALPEFTDVNQMYDALLNKGATGVTFDF-----PDTGVFLKGKISMDGGKAGVA 374
DB 215 --RQDILDLYHTQ-----GYPPDQNYTPGPGVRY-----PLTFGWCYKLV 255

QY 375 LVRSYKLYNKNSSNSTLKNLGEHRRARMDGGKAGVALVRSYKFYEDANGTRDHKKG 434
DB 256 PVEPD-KVEEANKGENTSLLHPVSLH---GMDDPEREVLWFRDSRL-----AF 300

QY 435 RHTARTSRDYKFDYDNKIDSTSGHHHHH 464
DB 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 5
US-10-687-060-15
; Sequence 15, Application US/10687060
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/EP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Pichia pastoris
```

```
US-10-687-060-15

Query Match      24.5%; Score 606; DB 17; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60

QY 79 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYVPN 138
DB 61 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYVPN 113

QY 139 RPLWKSHFRHTFEDEIEFIOGLEKSTGKVGIYPEIKAPWFHONGKIDIAAETLKVLL 198
DB 114 -----W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLE 147

QY 199 KYGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAIYTDWKETQKPKGYWV 258
DB 148 KHGAITSNTAATN-----AACAWLEAQEEBEVGPV 179

QY 259 NNYDWMFKPGMAEVKYGADVGPGWYMLVNKESKPDNIYVPLVKELAQYNVVHPY 318
DB 180 T-----PQVPLRPMTYKAADLSHPL---KEKGGLEGLHSQR----- 214

QY 319 TVRKDALPEFTDVNQMYDALLNKGATGVTFDF-----PDTGVFLKGKISMDGGKAGVA 374
DB 215 --RQDILDLYHTQ-----GYPPDQNYTPGPGVRY-----PLTFGWCYKLV 255

QY 375 LVRSYKLYNKNSSNSTLKNLGEHRRARMDGGKAGVALVRSYKFYEDANGTRDHKKG 434
DB 256 PVEPD-KVEEANKGENTSLLHPVSLH---GMDDPEREVLWFRDSRL-----AF 300

QY 435 RHTARTSRDYKFDYDNKIDSTSGHHHHH 464
DB 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 6
US-10-485-048-15
; Sequence 15, Application US/10485048
; Publication No. US20050058657A1
; GENERAL INFORMATION:
; APPLICANT: ERTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GPI20 and NEF and/or
; TITLE OF INVENTION: TAT for the Immunisation Against HIV
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-048-15

Query Match      24.5%; Score 606; DB 17; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60
```

Qy 79 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETKDGKQAVYVN 138
Db 61 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETMGK 113
Qy 139 RFLWKSHFRHTFEDEIEFIQGLEKSTGKVGIIYPEIKAPFHHQNGKQIAAETLKLVLK 198
Db 114 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
Qy 199 KYGDKKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEDPKGYWV 258
Db 148 KGAITSNTAATN-----AACAMLEAEQEEVEVGPV 179
Qy 259 NNYDWMFKPGAMAEVVKYADGVPGWYMLVKNKEBKPDNIYVYPLVKELAQYNVEVHPY 318
Db 180 T-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLIHSQR----- 214
Qy 319 TVRKDALPEFTFDVNMQYDALLNKSAGATGVFTDF-----PDTGVFELKGIKSMGKGAGVA 374
Db 215 --RQDILDWLVIYHTQ-----GYFPDQWNTYTPGPGVRY---PLTFGCWYKLV 255
Qy 375 LVRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKVFYEDANGTRDHKKG 434
Db 256 VPEPD-KVEEANKGENTSLHPVSLH---GMDDPREVLEWRFDLSRL-----AF 300
Qy 435 RHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
Db 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 7

US-10-203-013-19
; Sequence 19, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1el Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-19

Query Match 24.1%; Score 597; DB 14; Length 326;
Best Local Similarity 37.2%; Pred. No. 3.8e-40;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Qy 20 SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63
Qy 80 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETKDGKQAVYVN 139
Db 64 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETMGK 115
Qy 140 FPLWKSHFRHTFEDEIEFIQGLEKSTGKVGIIYPEIKAPFHHQNGKQIAAETLKLVLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 150
Qy 200 YGDKKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEDPKGYWV 259
Db 180 T-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLIHSQR----- 214
Qy 319 TVRKDALPEFTFDVNMQYDALLNKSAGATGVFTDF-----PDTGVFELKGIKSMGKGAGVA 375
Db 256 VPEPD-KVEEANKGENTSLHPVSLH---GMDDPREVLEWRFDLSRL-----AF 300
Qy 435 RHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
Db 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

Db 151 HGAITSNTAATN-----AACAMLEAEQEEVEVGPV 182
Qy 260 YNYDWMFKPGAMAEVVKYADGVPGWYMLVKNKEBKPDNIYVYPLVKELAQYNVEVHPY 319
Db 183 -----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLIHSQR----- 216
Qy 320 VRKDALPEFTFDVNMQYDALLNKSAGATGVFTDF-----PDTGVFELKGIKSMGKGAGVAL 375
Db 217 -RQDILDWLVIYHTQ-----GYFPDQWNTYTPGPGVRY---PLTFGCWYKLV 258
Qy 376 VRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKVFYEDANGTRDHKKGR 435
Db 259 VPEPD-KVEEANKGENTSLHPVSLH---GMDDPREVLEWRFDLSRL-----AFH 303
Qy 436 HTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

RESULT 8

US-10-687-060-19
; Sequence 19, Application US/10687060
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/EP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Pichia pastoris
US-10-687-060-19

Query Match 24.1%; Score 597; DB 17; Length 326;
Best Local Similarity 37.2%; Pred. No. 3.8e-40;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Qy 20 SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63
Qy 80 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETKDGKQAVYVN 139
Db 64 VVIHDFLDGLTDVAKFPFHRKDGRIYVYIDFTLKEIQSLEMTENFETMGK 115
Qy 140 FPLWKSHFRHTFEDEIEFIQGLEKSTGKVGIIYPEIKAPFHHQNGKQIAAETLKLVLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 150
Qy 200 YGDKKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEDPKGYWV 259
Db 151 HGAITSNTAATN-----AACAMLEAEQEEVEVGPV 182
Qy 260 YNYDWMFKPGAMAEVVKYADGVPGWYMLVKNKEBKPDNIYVYPLVKELAQYNVEVHPY 319
Db 183 -----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLIHSQR----- 216
Qy 320 VRKDALPEFTFDVNMQYDALLNKSAGATGVFTDF-----PDTGVFELKGIKSMGKGAGVAL 375
Db 217 -RQDILDWLVIYHTQ-----GYFPDQWNTYTPGPGVRY---PLTFGCWYKLV 258

QY 376 VRSDYKLYNKNSSNSTLKNLGEHHRARAMDGGKAGVALVRSDYKPYEDANGTRDHKKGR 435
Db 259 VEPPD-KVEEANKGENTSLHPSVLSH---GMDDPEREVLEWRFSRL-----AFH 303
QY 436 HTARTSRSDYKPYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

RESULT 9
US-10-485-048-19
; Sequence 19, Application US/10485048
; Publication No. US20050058657A1
; GENERAL INFORMATION:
; APPLICANT: ERTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
; TITLE OF INVENTION: TAT for the Immunisation Against HIV
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-485-048-19

Query Match 24.1%; Score 597; DB 17; Length 326;
Best Local Similarity 37.2%; Pred. No. 3.8e-40;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;
QY 20 SSSSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFACQADYLEQDLAMTKDGLV 79
Db 4 SSSSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFACQADYLEQDLAMTKDGLV 63
QY 80 VIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLENTENFETKQKQAVPNR 139
Db 64 VIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLENTENFETMGK-----115
QY 140 FPLWKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHONGKDIAAETLKVLLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 YGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDMKTEOKDPKGYWN 259
Db 151 HGATSSNTAATN-----AACAWLEAQSEEEVGFVPT 182
QY 260 YNYDMFKPGMAEVYKADGVGPWYMLNKEESKPDNIYVTPLVKELAQYNNVHPT 319
Db 183 -----PQVPLRPMTYKAADVLSHFL---KEKGGLEGLHSQR-----216
QY 320 VRKDALPEFTDQVNMVDALLNKSGATGVFTDF-----PDTGVEFLKGIKSMGDKAGVAL 375
Db 217 -RQDILDLMIVHTQ-----GYFPDQNYTPGPGVY-----PLTFGCYKLV 258
QY 376 VRSDYKLYNKNSSNSTLKNLGEHHRARAMDGGKAGVALVRSDYKPYEDANGTRDHKKGR 435
Db 259 VEPPD-KVEEANKGENTSLHPSVLSH---GMDDPEREVLEWRFSRL-----AFH 303
QY 436 HTARTSRSDYKPYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

RESULT 10
US-10-203-013-17
; Sequence 17, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1e1 Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-17

Query Match 24.0%; Score 594.5; DB 14; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
QY 19 CSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFACQADYLEQDLAMTKDGLR 78
Db 1 CSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFACQADYLEQDLAMTKDGLR 60
QY 79 VIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLENTENFETKQKQAVPN 138
Db 61 VIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLENTENFETMGK-----113
QY 139 RFPWLKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHONGKDIAAETLKVLLK 198
Db 114 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
QY 199 KYGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLK-LVOLIAYTDMK-----246
Db 148 KHGAITSNTAATNAACAWLEAQSEEEVGFVPTQVPLRPMTYKAAVDLSHFLKEKGGLE 207
QY 247 ---ETQEKDPKGYWV-----NINYDMFKPGMAEVYKADGVCP-----GW-YMLVN-290
Db 208 GLIHSQRQDILDLMIVHTQGYFPDQW-----NYTPGPGVRYPLTFGCYKLV 257
QY 291 -----KEESKPDNIYVTPLVKELA-----QYNVEVHPYTVRKDALPEFTD 331
Db 258 EPDKVEEANKGEN---TSLHPVLSLHGMDDEPEREVLEWRFSRLAFPHVARELHPEYFN 314
QY 332 VNQMYDALLNKSGATGVFTDFP-DTGVEFLKGIKSMGDKAGVALVRSDYKLYNKNSSN 390
Db 315 C-----TSEPVDPLRPWKHPCGSP-----KTACTN 340
QY 391 STLKVLGEH-----HPRAMDGGKAGVALVRSDYKPYEDANGTRDHKKGHRTARTSRD 444
Db 341 CYCKKCCFHCQVCFITKALGISYGR-----KKRRORRRPPQGSQTHQVLSLKOPTSQR 394
QY 445 YKFYDNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 11
US-10-687-060-17
; Sequence 17, Application US/10687060
; Publication No. US2005003022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine

APPLICANT: Godart, Stephane Andre Georges
APPLICANT: Marc-Hand, Martine
TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
TITLE OF INVENTION: and/or Nef Proteins
FILE REFERENCE: B45110
CURRENT APPLICATION NUMBER: US/10/687,060
CURRENT FILING DATE: 2003-10-16
PRIOR APPLICATION NUMBER: US/09/509,239
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/EP98/06040
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: GB 9720585.0
PRIOR FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 411
TYPE: PRT
ORGANISM: Pichia pastoris
US-10-687-060-17

Query Match 24.0%; Score 594.5; DB 17; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
Qy 19 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 78
Db 1 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60
Qy 79 VVIHDFLDGLTDVAKFPFHRHRKDGRIYVIDFTLKEIQSLEMTENFETDKGQAQVYPN 138
Db 61 VVIHDFLDGLTDVAKFPFHRHRKDGRIYVIDFTLKEIQSLEMTENFETDKGQAQVYPN 113
Qy 139 RPLWKSHFRHITFEDETEFTQGLEKSTGKVGIYPEIKAPWFHONGKDIATLKV 198
Db 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
Qy 199 KGVYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK--- 246
Db 148 KHGAITSNTAATNAACAWLEAQEEVGFVPTQVPLRPMYKAAVDLSHFLKEKGGLE 207
Qy 247 ----ETQKDPKGYVW-----NINYDMMPKPGAMAEVVKYADGVGP-----GW-YMLVN- 290
Db 208 GLIHSQRQDILDWIYHTQGYFDMQ-----NYTPGPGVRYPLTFGWCYKLVVP 257
Qy 291 -----KEESKPDNIVYPLVKELA-----QYNVEVHPYTVRKDALPEPFTD 331
Db 258 EPDKVEEANKGEN---TSLHPVSLHGMDDPEREVLWRFSRLAFHHVARELHPEYFKN 314
Qy 332 VNQMYDALLNKGATGVPTDPP-DTGVEFLKGIKMDGGKAGVALVRSYKLYNKNSSN 390
Db 315 C-----TSEPVDPRLEPWKPGSQP-----KTACTN 340
Qy 391 STLKNLGEH-----HRAAMDGGKAGVALVRSYKLYNKNSSN 444
Db 341 CYCKKCCFHCQVCFTKALGISYGR-----KKRRQRRPPQGSQTHQVSLSKOPTSQSR 394
Qy 445 YKFDVKNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 12

US-10-485-048-17
Sequence 17, Application US/10485048
Publication No. US20050058657A1
GENERAL INFORMATION:
APPLICANT: ERTL, Peter Franz
APPLICANT: TITE, John Philip
APPLICANT: VAN WELY, Catherine Anne
APPLICANT: VOSS, Gerald
TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
TITLE OF INVENTION: TAT for the Immunisation Against HIV

FILE REFERENCE: B45279
CURRENT APPLICATION NUMBER: US/10/485,048
CURRENT FILING DATE: 2004-01-26
PRIOR APPLICATION NUMBER: PCT/EP02/08343
PRIOR FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: GB 0118367.2
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 98
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-10-485-048-17

Query Match 24.0%; Score 594.5; DB 17; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
Qy 19 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 78
Db 1 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60
Qy 79 VVIHDFLDGLTDVAKFPFHRHRKDGRIYVIDFTLKEIQSLEMTENFETDKGQAQVYPN 138
Db 61 VVIHDFLDGLTDVAKFPFHRHRKDGRIYVIDFTLKEIQSLEMTENFETDKGQAQVYPN 113
Qy 139 RPLWKSHFRHITFEDETEFTQGLEKSTGKVGIYPEIKAPWFHONGKDIATLKV 198
Db 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
Qy 199 KGVYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK--- 246
Db 148 KHGAITSNTAATNAACAWLEAQEEVGFVPTQVPLRPMYKAAVDLSHFLKEKGGLE 207
Qy 247 ----ETQKDPKGYVW-----NINYDMMPKPGAMAEVVKYADGVGP-----GW-YMLVN- 290
Db 208 GLIHSQRQDILDWIYHTQGYFDMQ-----NYTPGPGVRYPLTFGWCYKLVVP 257
Qy 291 -----KEESKPDNIVYPLVKELA-----QYNVEVHPYTVRKDALPEPFTD 331
Db 258 EPDKVEEANKGEN---TSLHPVSLHGMDDPEREVLWRFSRLAFHHVARELHPEYFKN 314
Qy 332 VNQMYDALLNKGATGVPTDPP-DTGVEFLKGIKMDGGKAGVALVRSYKLYNKNSSN 390
Db 315 C-----TSEPVDPRLEPWKPGSQP-----KTACTN 340
Qy 391 STLKNLGEH-----HRAAMDGGKAGVALVRSYKLYNKNSSN 444
Db 341 CYCKKCCFHCQVCFTKALGISYGR-----KKRRQRRPPQGSQTHQVSLSKOPTSQSR 394
Qy 445 YKFDVKNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 13

US-10-203-013-21
Sequence 21, Application US/10203013
Publication No. US20030158134A1
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham Biologicals S.A.
TITLE OF INVENTION: No. US20030158134A1el Use
FILE REFERENCE: B45209
CURRENT APPLICATION NUMBER: US/10/203,013
CURRENT FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: GB 0002200.4
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: GB 0009336.9
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: GB 0013806.5
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: PCT/EP00/05998

```
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-21

Query Match      23.6%; Score 585.5; DB 14; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

QY 20 SSSSNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLBODLAWTKDGRIV 79
Db 4 SSSSNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLBODLAWTKDGRIV 63
QY 80 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETKDGKQAQVYNR 139
Db 64 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETMGGK----- 115
QY 140 FPLKSHRIHTFEDEIEFIQGLEKSTGKKGVIYPEIKAPWPHQNGKIDIAAETLKLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 GYDVKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246
Db 151 HGAITSSNTAATNAACAWLEAQEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEGGLEG 210
QY 247 ---ETQEKDPKGYV---NYYDMFKPGMAEYVVKYADGVGP-----GW-YMLVN-- 290
Db 211 LIHSQRQDILDLMYHTQGYFPDWQ-----NYTPGFGVRYPLTFGMCYKLVPE 260
QY 291 ---KEESKPDNIYVTVPLVKELA-----OYNVEVHPYTVRKDALPEFTDV 332
Db 261 PDKVEEANKGEN---TSLHPVSLHGMDDPEREVLEWFRDLSLAFHHVARELHPEYFNC 317
QY 333 NQMDYDALLNKSGATGVFTDFP-DTGVEFLKGIKSGMDGKAGVALVRSYKLYNKNSSNS 391
Db 318 -----TSEPVDPRLEPKWHPGSGP-----KTACTNC 343
QY 392 TLKNLGEH-----HRRAMDGGKAGVALVRSYDYKFYEDANGTRDHKKGRHTARTSRDY 445
Db 344 YCKKCCFHCQVCFTIKALGISYGR-----KKRRQRRPPQSGSOTHQVSLSKQPTSQSR- 396
QY 446 KFYDNKRIDSTSGHHHHH 464
Db 397 --GDPTGPKETSGHHHHH 413

RESULT 14
US-10-687-060-21
; Sequence 21, Application US/10687060
; Publication No. US2005003022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/EP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 413

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-048-21

Query Match      23.6%; Score 585.5; DB 17; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

QY 20 SSSSNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLBODLAWTKDGRIV 79
Db 4 SSSSNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLBODLAWTKDGRIV 63
QY 80 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETKDGKQAQVYNR 139
Db 64 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETMGGK----- 115
QY 140 FPLKSHRIHTFEDEIEFIQGLEKSTGKKGVIYPEIKAPWPHQNGKIDIAAETLKLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 GYDVKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246
Db 151 HGAITSSNTAATNAACAWLEAQEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEGGLEG 210
QY 247 ---ETQEKDPKGYV---NYYDMFKPGMAEYVVKYADGVGP-----GW-YMLVN-- 290
Db 211 LIHSQRQDILDLMYHTQGYFPDWQ-----NYTPGFGVRYPLTFGMCYKLVPE 260
QY 291 ---KEESKPDNIYVTVPLVKELA-----OYNVEVHPYTVRKDALPEFTDV 332
Db 261 PDKVEEANKGEN---TSLHPVSLHGMDDPEREVLEWFRDLSLAFHHVARELHPEYFNC 317
QY 333 NQMDYDALLNKSGATGVFTDFP-DTGVEFLKGIKSGMDGKAGVALVRSYKLYNKNSSNS 391
Db 318 -----TSEPVDPRLEPKWHPGSGP-----KTACTNC 343
QY 392 TLKNLGEH-----HRRAMDGGKAGVALVRSYDYKFYEDANGTRDHKKGRHTARTSRDY 445
Db 344 YCKKCCFHCQVCFTIKALGISYGR-----KKRRQRRPPQSGSOTHQVSLSKQPTSQSR- 396
QY 446 KFYDNKRIDSTSGHHHHH 464
Db 397 --GDPTGPKETSGHHHHH 413

RESULT 15
US-10-485-048-21
; Sequence 21, Application US/10485048
; Publication No. US20050058657A1
; GENERAL INFORMATION:
; APPLICANT: BRTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GPI20 and NEF and/or
; TITLE OF INVENTION: TAT for the Immunisation Against HIV
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-485-048-21

Query Match      23.6%; Score 585.5; DB 17; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
```

Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Qy 20 SSSSNMANTOMKSDKIIIAHRGAGYLPEHTLESKALAFQAQADYLEODLAMTKDGLV 79
Db |||||
Qy 4 SSSSNMANTOMKSDKIIIAHRGAGYLPEHTLESKALAFQAQADYLEODLAMTKDGLV 63
Db |||||
Qy 80 VIHDFLDGLTDVAKKFPHRHKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYVNR 139
Db |||||
Qy 64 VIHDFLDGLTDVAKKFPHRHKDGRYYVIDFTLKEIQSLEMTENFETMGCK----- 115
Db |||||
Qy 140 FPLMKSHPRINTFEDEIBFIQGLEKSTGKKGVIPEIKAPWFHQNGKDIAAETLKVLLK 199
Db |||||
Qy 116 ---W-----SKSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
Db |||||
Qy 200 YGYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246
Db |||||
Qy 151 HGAITSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKCKGGLG 210
Db |||||
Qy 247 ---ETQEKDPKGYWV---NINYDMFKPGAMAEVVKYADGVGP-----GW-YMLVN-- 290
Db |||||
Qy 211 LIHSORQDILDLMYHTQGYFPDQ-----NYTPGVRYPLTFGCYKLVPE 260
Db |||||
Qy 291 ----KEESKPDNIYVTVPLVKELA-----QYNVEVHPYTVRKDALPEFFTDV 332
Db |||||
Qy 261 PDKVEEANKGEN---TSLHPVSLHGMDDPEREVLEWRFSRLAFHHVARELHPEYFKNC 317
Db |||||
Qy 333 NQMYDALLNKSGATGVFTDFP-DTCVEFLKGIKSMGDKGAGVALVRSYKLYNKNSSSNS 391
Db |||||
Qy 318 -----TSEPVDPRLEPWKHGSGQP-----KTACTNC 343
Db |||||
Qy 392 TLKNLGEH-----HRAMMDGGKAGVALVRSYKFYEDANGTRDHKKGRHTARTSRSDY 445
Db |||||
Qy 344 YCKKCCFHCQVCFITKALGISYGR-----KKRQRPRPQGSQTHQVLSKQPTSQSR- 396
Db |||||
Qy 446 KFYDNKRIDSTSGHHHHH 464
Db --GDPTGPKETSGHHHHH 413

Search completed: June 22, 2005, 11:54:14
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:31:10 ; Search time 41 Seconds
(without alignments)
1088.892 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLALSLLAAGVLGCS.....YKFYDKRIDSTSGHHHHH 464
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1920	77.5	364	2 A43576	glycerophosphodies
2	1916	77.4	364	2 S59932	glycerophosphodies
3	1914	77.3	364	2 S59931	glycerophosphodies
4	1910	77.1	364	2 S59934	glycerophosphodies
5	1902	76.8	364	2 G64086	glycerophosphodies
6	1885	76.1	364	2 S59933	glycerophosphodies
7	1885	76.1	364	2 S59936	glycerophosphodies
8	1261.5	50.9	358	2 D91019	glycerophosphodies
9	1261.5	50.9	358	2 F85863	glycerophosphodies
10	1251.5	50.5	358	2 S15945	glycerophosphodies
11	1233	49.8	356	2 AH0791	glycerophosphodies
12	1210	48.8	371	2 AC0466	glycerophosphodies
13	1074	43.4	356	2 F71346	probable glyceroph
14	957	38.6	371	2 H82497	glycerophosphoryl
15	425	17.2	1027	2 AC1841	glycerophosphoryl
16	415.5	16.8	786	2 D75630	glycerophosphoryl
17	383	15.5	383	2 A83603	glycerophosphoryl
18	368	14.9	356	2 F87654	glycerophosphoryl
19	354	14.3	389	2 AH1937	glycerophosphoryl
20	353	14.3	372	2 T45628	glycerophosphodies
21	347.5	14.0	392	2 D96770	hypothetical prote
22	321.5	13.0	293	2 I04018	glycerophosphoryl
23	269.5	10.9	309	2 H89862	glycerophosphoryl
24	262.5	10.6	284	2 D84112	glycerophosphoryl
25	246.5	10.0	375	2 A83352	probable glyceroph
26	222	9.0	243	2 E69827	glycerophosphodies
27	181	7.3	247	2 E89956	hypothetical prote
28	179.5	7.2	239	2 F69961	glycerophosphodies
29	172.5	7.0	583	2 AH1151	glycerophosphoryl

30	171	6.9	249	2 G96952	glycerophosphoryl
31	170.5	6.9	583	2 A11510	C-terminal domain
32	168	6.8	232	2 AD1236	glycerophosphodies
33	167	6.7	235	2 A11598	glycerophosphodies
34	165.5	6.7	285	2 E75317	glycerophosphoryl
35	156.5	6.3	240	2 G90544	hypothetical prote
36	154	6.2	353	2 C64187	outer membrane pro
37	152.5	6.2	268	2 AF3245	agrocinoipine phosp
38	151.5	6.1	763	2 F96693	hypothetical prote
39	148.5	6.0	271	2 B86918	hypothetical prote
40	147.5	6.0	241	2 S73747	glycerophosphoryl
41	146	5.9	770	2 T04792	hypothetical prote
42	145	5.9	587	2 G89785	hypothetical prote
43	144.5	5.8	244	2 D84232	glycerophosphoryl
44	142.5	5.8	250	2 H83784	glycerophosphodies
45	142	5.7	281	2 H86632	conserved hypothet

ALIGNMENTS

RESULT 1

A43576
Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influenzae
N:Alternate names: immunoglobulin D-binding protein; protein D
C:Species: Haemophilus influenzae
A:Variety: isolate 772
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 08-Oct-1999
C:Accession: A43576
R:Janson, H.; Heden, L.O.; Grubb, A.; Ruan, M.; Forsgren, A.
Infect. Immun. 59, 119-125, 1991
A:Title: Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: clon
A:Reference number: A43576; MUID:91099948; PMID:1987023
A:Accession: A43576
A:Molecule type: DNA
A:Residues: 1-364 <JAN>
A:Cross-references: GB:M37487; NID:g148970; PIDN:AAA24998.1; PID:g148971
A:Experimental source: isolate 772
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-364/Product: glycerophosphodiester phosphodiesterase #status predicted <MAT>

Query Match 77.5%; Score 1920; DB 2; Length 364;
Best Local Similarity 99.5%; Pred. No. 1.4e-122;
Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MDPKTLALSLLAAGVLGCS	SHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAPA	60
Db	1	MKLKTLALSLLAAGVLGCS	SHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAPA	60
Qy	61	QOADYLEQDLAMTKDGR	LVIHDFLDGLTDVAKKFPFHRKDKGRYYVIDFTLKEIQSLE	120
Db	61	QOADYLEQDLAMTKDGR	LVIHDFLDGLTDVAKKFPFHRKDKGRYYVIDFTLKEIQSLE	120
Qy	121	MTEFETKDGKQAQVY	PNRFPFLWKSHPRIFTFEDIEIFIOGLEKSTGKVGVIPEIKAPW	180
Db	121	MTEFETKDGKQAQVY	PNRFPFLWKSHPRIFTFEDIEIFIOGLEKSTGKVGVIPEIKAPW	180
Qy	181	FHHQNGKDIAAETL	KVYGVYDKKTDVYLTQDPNELKRIKTELLPQMGMDLKLVLQI	240
Db	181	FHHQNGKDIAAETL	KVYGVYDKKTDVYLTQDPNELKRIKTELLPQMGMDLKLVLQI	240
Qy	241	AYTDWKETQEKDP	KGYWYNNYNDWMFKGMAEAVVYKADGVGPGWYMLVNKEESKPDNIV	300
Db	241	AYTDWKETQEKDP	KGYWYNNYNDWMFKGMAEAVVYKADGVGPGWYMLVNKEESKPDNIV	300
Qy	301	YTPLVKELAQYN	VEVHPYTVRKDALPEFTDQNMQYDALLNKS	360
Db	301	YTPLVKELAQYN	VEVHPYTVRKDALPEFTDQNMQYDALLNKS	360
Qy	361	KGIK	364	
Db	361	KGIK	364	


```
Db 121 MTENFETKDGKQAQVYNNRFLPWKSHFRHTDEIEFTQGLEKSTGKVGVIPEIKAPW 180
Qy 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Qy 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Qy 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Qy 361 KGIK 364
Db 361 KGIK 364

RESULT 5
G64086
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain R
N:Alternate names: Igb-binding protein; protein D
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: G64086
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: G64086
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-364 <TIGR>
A:Cross-references: UNIPROT:Q06282; GB:U32751; GB:L42023; NID:g3212200; PIDN:AAC22348.1;
C:Experimental source: strain Rd KW20
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 76.8%; Score 1902; DB 2; Length 364;
Best Local Similarity 98.4%; Pred. No. 2.4e-121;
Matches 358; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHRCASGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHRCASGYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKPPHRRKDGRIYVYIDFTLKEIQSLE 120
Db 61 QHSDYLEODLAWTKDGRVLIHDFLDGLTDVAKKPPHRRKDGRIYVYIDFTLKEIQSLE 120
Qy 121 MTENFETKDGKQAQVYNNRFLPWKSHFRHTDEIEFTQGLEKSTGKVGVIPEIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFLPWKSHFRHTDEIEFTQGLEKSTGKVGVIPEIKAPW 180
Qy 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Qy 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Qy 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Qy 361 KGIK 364
Db 361 KGIK 364
```

RESULT 6

```
S59933
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 6
N:Alternate names: immunoglobulin D-binding protein; protein D
C:Species: Haemophilus influenzae
A:Variety: strain 6-7626
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C:Accession: S59933; S47335
R:Song, X.M.; Forsgren, A.; Jansson, H.
Infect. Immun. 63, 696-699, 1995
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen
A:Reference number: S59931; MUID:95122210; PMID:7822043
A:Accession: S59933
A:Molecule type: DNA
A:Residues: 1-364 <SON>
A:Cross-references: EMBL:Z35658; NID:g525217; PIDN:CAA84717.1; PID:g525218
C:Experimental source: strain 6-7626
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 76.1%; Score 1885; DB 2; Length 364;
Best Local Similarity 97.5%; Pred. No. 3.4e-120;
Matches 355; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHRCASGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHRCASGYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKPPHRRKDGRIYVYIDFTLKEIQSLE 120
Db 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKPPHRRKDGRIYVYIDFTLKEIQSLE 120
Qy 121 MTENFETKDGKQAQVYNNRFLPWKSHFRHTDEIEFTQGLEKSTGKVGVIPEIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFLPWKSHFRHTDEIEFTQGLEKSTGKVGVIPEIKAPW 180
Qy 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKYDYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Qy 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKETQEKDPKPGYNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Qy 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVTDPDGTGVBEFL 360
Qy 361 KGIK 364
Db 361 KGIE 364

RESULT 7
S59936
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N
N:Alternate names: Igb-binding protein; protein D
C:Species: Haemophilus influenzae
A:Variety: strain NCTC 8468
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C:Accession: S59936; S47338
R:Song, X.M.; Forsgren, A.; Jansson, H.
Infect. Immun. 63, 696-699, 1995
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen
A:Reference number: S59931; MUID:95122210; PMID:7822043
A:Accession: S59936
A:Molecule type: DNA
A:Residues: 1-364 <SON>
A:Cross-references: EMBL:Z35661; NID:g525223; PIDN:CAA84720.1; PID:g525224
C:Experimental source: strain NCTC 8468
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen
```

Query Match 76.1%; Score 1885; DB 2; Length 364;
Best Local Similarity 97.5%; Pred. No. 3.4e-120;
Matches 355; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHKGAGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLATGTVAGCSSHSSNMANTQMSKIIIAHKGAGYLPEHTLESKALAPA 60

Qy 61 QOADLEODLAWTKDGLRWLVHDFDLGLTDVAKKFPHRHDKGRYYVIDFTLKEIOSLE 120
Db 61 QHADYLEODLAWTKDGLRWLVHDFDLGLTDVAKKFPHRHDKGRYYVIDFTLKEIOSLE 120

Qy 121 MTENFETKDGKQAQVYNNRFPPLWKSHPRIHFEDEIEFIQGLEKSTGKKGVIYEIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFPPLWKSHPRIHFEDEIEFIQGLEKSTGKKGVIYEIKAPW 180

Qy 181 FHHQNGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240

Qy 241 AYTDWKETOEKDPKGYWVNNYNDWMPKPGAAEVVYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKETOEKDPKGYWVNNYNDWMPKPGAAEVVYADGVGPGWYMLVNKEESKPDNIV 300

Qy 301 YTPLVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKGATGVFTDFPDGTGVEFL 360
Db 301 YTPLVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKGATGVFTDFPDGTGVEFL 360

Qy 361 KGIK 364
Db 361 KGIK 364

RESULT 8
D91019
glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, #
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D91019
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D91019
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <HAY>
A:Cross-references: UNIPROT:Q8XE17; GB:BA000007; PIDN:BA036547.1; PID:gl3362594; GSPDB:C
A:Experimental source: strain O157:H7, substrain RIND 0509952
C:Genetics:
A:Gene: EC9124

Query Match 50.9%; Score 1261.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 5.3e-78;
Matches 231; Conservative 55; Mismatches 64; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHKGAGYLPEHTLESKALAPAQA 63
Db 6 KNLSSMAIMSTIVMGSSAMAAD-----SNEKIVIAHKGAGYLPEHTLPKAWAYAQGA 59

Qy 64 DYLEDODLAWTKDGLRWLVHDFDLGLTDVAKKFPHRHDKGRYYVIDFTLKEIOSLEWTE 123
Db 60 DYLEDQDLVMTKDDHLVHLDHVRDVAADFPDRARKGRYYAIDFTLDEIKSLKETE 119

Qy 124 NFKETKDGKQAQVYNNRFPPLWKSHPRIHFEDEIEFIQGLEKSTGKKGVIYEIKAPWPHH 183
Db 120 GFDIENGKKVQYPCGRFPMGKSDFRVHTFESEIEFVQGLNHSHTGKNIIGIYEIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 180 QEGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 239

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAAEVVYADGVGPGWYMLVNKEESKPDNIVVYTP 303
Db 240 DWNETQOKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTDVNQLYDALLNKGAVNGLFTDFDPKAVKFL 355

RESULT 10
S15945
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich
C:Species: Escherichia coli
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: S15945; S00871; E64994; S72653; S14522
R:Tommassen, J.; Biglmeier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.
Mol. Gen. Genet. 226, 321-327, 1991
A:Title: Characterization of two genes, glpQ and ugpQ, encoding glycerophosphoryl dieste
A:Reference number: S15945; MUID:91238712; PMID:1851953
A:Accession: S15945
A:Molecule type: DNA
A:Residues: 1-358 <TOM>
A:Cross-references: UNIPROT:P09394; EMBL:X56907; NID:g41580; PIDN:CAA40223.1; PID:g41581
R:Eiglmeier, K.; Boos, W.; Cole, S.T.

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAAEVVYADGVGPGWYMLVNKEESKPDNIVVYTP 303
Db 240 DWNETQOKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTDVNQLYDALLNKGAVNGLFTDFDPKAVKFL 355

RESULT 9
F85863
glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, #
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85863
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85863
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STO>
A:Cross-references: UNIPROT:Q8XE17; GB:AB005174; NID:gl2516579; PIDN:AAG57370.1; GSPDB:C
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: glpQ

Query Match 50.9%; Score 1261.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 5.3e-78;
Matches 231; Conservative 55; Mismatches 64; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLAGCSSHSSNMANTQMSKIIIAHKGAGYLPEHTLESKALAPAQA 63
Db 6 KNLSSMAIMSTIVMGSSAMAAD-----SNEKIVIAHKGAGYLPEHTLPKAWAYAQGA 59

Qy 64 DYLEDODLAWTKDGLRWLVHDFDLGLTDVAKKFPHRHDKGRYYVIDFTLKEIOSLEWTE 123
Db 60 DYLEDQDLVMTKDDHLVHLDHVRDVAADFPDRARKGRYYAIDFTLDEIKSLKETE 119

Qy 124 NFKETKDGKQAQVYNNRFPPLWKSHPRIHFEDEIEFIQGLEKSTGKKGVIYEIKAPWPHH 183
Db 120 GFDIENGKKVQYPCGRFPMGKSDFRVHTFESEIEFVQGLNHSHTGKNIIGIYEIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 180 QEGKDIAAETLKVLYKKGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 239

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAAEVVYADGVGPGWYMLVNKEESKPDNIVVYTP 303
Db 240 DWNETQOKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTDVNQLYDALLNKGAVNGLFTDFDPKAVKFL 355

RESULT 10
S15945
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich
C:Species: Escherichia coli
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: S15945; S00871; E64994; S72653; S14522
R:Tommassen, J.; Biglmeier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.
Mol. Gen. Genet. 226, 321-327, 1991
A:Title: Characterization of two genes, glpQ and ugpQ, encoding glycerophosphoryl dieste
A:Reference number: S15945; MUID:91238712; PMID:1851953
A:Accession: S15945
A:Molecule type: DNA
A:Residues: 1-358 <TOM>
A:Cross-references: UNIPROT:P09394; EMBL:X56907; NID:g41580; PIDN:CAA40223.1; PID:g41581
R:Eiglmeier, K.; Boos, W.; Cole, S.T.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AH0791

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-356 <PAR>

A;Cross-references: GB:ALU513382; PIDN:CAD07514.1; PID:g16503507; GSPDB:GN00176

C;Genetics:

A;Gene: STY2511

C;Keywords: phosphoric diester hydrolase

Query Match 49.8%; Score 1233; DB 2; Length 356;
Best Local Similarity 64.5%; Pred.No.4.5e-76;
Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;

Qy 4 KTLALSLLAAGVLACGSSSHSNMANTQMKS-----VAAEKVVIAHRAASGYLPEHTLESKALAFQAQA 63
Db 6 KNLSVALMLAGMTTIGSGA-----VAAEKVVIAHRAASGYLPEHTLPKAMAYAQGA 56

Qy 64 DYLEQDLAMTKDGLRLVTHDFLDGLTDVAKKFPHRRKDGRYYVIDFTLKEIQSLEWTE 123
Db 57 DYLEQDLVMTKDDHLVLDHYLDRTVDADFPDRARKDGRIYAIDFTLDEIKSLKTFE 116

Qy 124 NFETKDKQAQVYNRPFPLWKSHFRHIHPEDIEFTQGLESTCKKVGIYPIEIKAPWFHH 183
Db 117 GFDIENGKKVQTYPGRFPMGSKSDFIHTFEIEFVQGLNHSTGKNIGIYPIEIKAPWFHH 176

Qy 184 QNGKXIDIAAEITLVLKCYGDYKKTDMVYLQTFDNELKRITKELLPOGMMDLKLVLIIAYT 243
Db 177 QEGKXIDIAAEITLVLKCYGTIGKQDNVYLQCFDVAELEKRIKNELEPRKMGMNLNVOLIAYT 236

Qy 244 DWKETQBDPKGYVNNYNDWMFKPAMAEEVKYADGVGPWGMYLVNKESKPDNIIVYTP 303
Db 237 DWNETQQKQDPGRVNNYNDWMFKPGAMKQVAEYADGIGDPYHMLV-AEGSTRGNIKLTG 295

Qy 304 LVKELAQNVEVHPYTVRKDALPEFTFDVNQMYDALLNKSGATGVTTDFPDGTGVEFLK 361
Db 296 MVQDAHQNMVHPYTVRADQLPDYATDVNQLYDYILYNKAGVDGLFTDFPDKAVMFLO 353

RESULT 12

AC0466

Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) [imported] - Yersinia pestis (stra

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C;Accession: AC0466

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Ruthersford, K.; Simmonds, T.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AC0466

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-371 <KUR>

A;Cross-references: UNIPROT:Q8ZAH4; GB:AL590842; PIDN:CAC93295.1; PID:g15981742; GSPDB:G

C;Genetics:

A;Gene: gfpQ

C;Keywords: phosphoric diester hydrolase

Query Match 48.8%; Score 1210; DB 2; Length 371;
Best Local Similarity 62.7%; Pred.No.1.7e-74;
Matches 227; Conservative 54; Mismatches 75; Indels 6; Gaps 4;

Qy 4 KTLALSLLAAGVLACGSSSHSNMANTQMKS---DKIIIAHRAASGYLPEHTLESKALAPA 60
Db 6 KTLASLILAASMA-APAYAAPKASADKSASAIKQVIAHRAASGYLPEHSLPAKAMAYA 64

Qy 61 QQADYLEQDLAMTKDGLRLVTHDFLDGLTDVAKKFPHRRKDGRYYVIDFTLKEIQSLE 120
Db 65 QGADYLEQDLVMTKDDHLVLDHYLDRTVDADFPDRARKDGRIYAIDFTLPEIKSLK 124


```
Query Match      17.2%; Score 425; DB 2; Length 1027;
Best Local Similarity   29.4%; Pred.No. 9.6e-21;
Matches 137; Conservative 66; Mismatches 147; Indels 116; Gaps 20;
```

Qy	37	I I A H R G S G Y L P E H T L S K A L A F A Q O A D Y L E O D L A M T K D G L V I H D - - - - -	H 84
Db	180	V I G H R G A S G R P E H T L S Y K U A I B Q G A D F I E P D L A V T K D G V L I A R H E P A L A V L N A D G S V N	239
Qy	85	F L O G L T O V A K - - K P P H R R K - - - D G R - - - Y Y V I D F T L K E I Q S L E M T E N P E T K D G K Q A Q V	135
Db	240	F S N T T T V Y Q I A K E S D R L K T V N L D G T R I T G W F A E D F T L A E I K E L R A J E R L - - - - -	289
Qy	136	Y P N R F P L W K G H F R I H T P E D E I E F I Q G L E K S T G K K V G I Y P E I K A P W F H Q - - - - -	184
Db	290	- P P R Q S F N G Q F T P T L A E I I D L V K Q V E A T G K K I G I Y P E T K H T Y P A Q E A T Y V G T T E K I	348
Qy	185	-- N G K D T A A E T L K V L K Y G V D K K T D -- M V Y L Q T P D F N E L K R I K T E L L P Q W G M D L K I V O L I	240
Db	349	N R N I S Q I L D I T L K A - - - - - N N F T D P S R I F T Q S P E V G N L K E L H D T I M P N A G V D I P L V Q L F	402
Qy	241	A Y T D W K E T Q E K D P K Y G V W Y N Y - Y D W M F K - - - - - P G A M E A V K Y A D G V C P G W Y M L	288
Db	403	- - - - - D A I D V D I N G R L I E T R P D Y F I V S G D T R T Y G D L R T P A G L A E I A E Y A D G I G P W K R M I	456
Qy	289	V N K E E S K P D N - - - - - I V Y T P L V K E L A Q Y N V E V H P Y T V R K D A L P E F	328
Db	457	V S V R G T D A N D G Q A D V N G D G A V N D A K T L L P P T T L V Q D A H N V G L Q V H P Y T F R D E - - - - -	513
Qy	329	F T D V N Q M - - - - - Y D A L L N K S G A T G V F T D P D T G V E F L K G I K S M D G G K A G V A L V R S D Y K L	382
Db	514	Y L A A N Y Q G N P E L E Y Q Q L F - Q L G V D A L F T D F P I T - A D R V R D R L S L P G N N - - - - - I V R S P - - -	564
Qy	383	Y N K N S S N S T L K N L G E H H R A M D G G K A G V A L V R S D Y K Y F E D A N G T	428
Db	565	Q N P D V L S G D A F A N I G - - - G S R G F E G G - - - A I N A S K T K L Y M L E G T	603

Search completed: June 22, 2005, 11:41:29
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:30:25 ; Search time 171 Seconds
(without alignments)
1389.501 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLALSLAAGVLACGS.....YKFYDNKRIDSTGHHHHH 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_prot.*

2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1902	76.8	364	1 GLPQ_HAEIN	Q06282 haemophilus
2	1577	63.7	358	2 Q9CLO4	Q9CLO4 pasteurella
3	1501.5	60.6	368	2 Q65R11	Q65R11 manheimia
4	1498.5	60.5	357	2 Q8RHT6	Q8RHT6 fusobacteri
5	1363	55.0	309	2 Q7P6W6	Q7P6W6 fusobacteri
6	1261.5	50.9	358	2 Q8XE17	Q8XE17 escherichia
7	1254.5	50.6	358	2 Q8CVV6	Q8CVV6 escherichia
8	1251.5	50.5	358	1 GLPQ_ECOLI	P09394 escherichia
9	1251.5	50.5	358	2 Q7UC72	Q7UC72 shigella fl
10	1251.5	50.5	377	2 Q83QU2	Q83QU2 shigella fl
11	1248	50.4	361	2 Q6CZ14	Q6CZ14 erwinia car
12	1235	49.9	356	2 Q8ZNG7	Q8ZNG7 salmonella
13	1233	49.8	356	2 Q8Z556	Q8Z556 salmonella
14	1210.5	48.9	358	2 Q7MZY7	Q7MZY7 photorhabdu
15	1210.5	48.8	371	2 Q66FX7	Q66FX7 yersinia ps
16	1210	48.8	371	2 Q8ZAH4	Q8ZAH4 yersinia pe
17	1210	48.8	372	2 Q8D1J5	Q8D1J5 yersinia pe
18	1082	43.7	356	2 Q9X5U6	Q9X5U6 treponema p
19	1074	43.4	356	1 GLPQ_TREPA	Q30405 treponema p
20	1074	43.4	356	2 Q71U57	Q71U57 treponema s
21	1074	43.4	356	2 Q71U58	Q71U58 treponema p
22	1074	43.4	356	2 Q71U59	Q71U59 treponema p
23	1074	43.4	356	2 Q71U64	Q71U64 treponema p
24	1072.5	43.3	341	2 Q7QMYS	Q7QMYS anopheles g
25	1026	41.4	360	2 Q6LT65	Q6LT65 photobacter
26	987.5	39.9	351	2 Q87M77	Q87M77 vibrio para
27	979	39.5	352	2 Q7MDR9	Q7MDR9 vibrio vuln
28	979	39.5	352	2 Q8D6W7	Q8D6W7 vibrio vuln
29	957	38.6	371	2 Q9KN30	Q9KN30 vibrio chol
30	839.5	33.9	338	2 Q9F8D4	Q9F8D4 borrelia tu
31	836	33.8	337	2 Q9F8D3	Q9F8D3 borrelia co

32	830.5	33.5	338	2	Q9F8D5	Q9F8D5 borrelia pa
33	825.5	33.3	334	2	Q6UKS8	Q6UKS8 borrelia mi
34	821.5	33.2	340	2	Q45201	Q45201 borrelia he
35	811.5	32.8	336	2	Q6UKS9	Q6UKS9 borrelia lo
36	803.5	32.4	333	2	Q9F8D6	Q9F8D6 borrelia cr
37	797.5	32.2	333	2	Q9ETK6	Q9ETK6 borrelia re
38	426.5	17.3	372	2	Q7YZB5	Q7YZB5 bombyx mori
39	425	17.2	1027	2	Q8Z028	Q8Z028 anabaena sp
40	415.5	16.8	786	2	Q9RZL4	Q9RZL4 deinococcus
41	385	15.5	392	2	Q8CJ18	Q8CJ18 streptomyce
42	383	15.5	383	2	Q916E6	Q916E6 pseudomonas
43	368	14.9	356	2	Q9A3D2	Q9A3D2 caulobacter
44	360	14.5	393	2	Q827Y5	Q827Y5 streptomyce
45	354	14.3	389	2	Q8Y104	Q8Y104 anabaena sp

ALIGNMENTS

RESULT 1

ID	GLPQ_HAEIN	STANDARD;	PRT;	364 AA.
AC	Q06282;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Glycerophosphoryl diester phosphodiesterase precursor (EC 3.1.4.46)			
DE	(Glycerophosphodiester phosphodiesterase) (Surface-exposed lipoprotein			
DE	D) (Protein D) (Immunoglobulin D-binding protein) (IgD-binding			
DE	protein).			
GN	Name=gipQ; Synonyms=hpd; OrderedLocusNames=HI0689;			
OS	Haemophilus influenzae.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Haemophilus.			
OX	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RD / KW20 / ATCC 51907;			
RC	MEDLINE=95350630; PubMed=7542800;			
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,			
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,			
RA	McKenny K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,			
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,			
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,			
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,			
RA	Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,			
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,			
RA	Venter J.C.;			
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae			
RT	Rd.";			
RL	Science 269:496-512(1995).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NTHi 772;			
RC	MEDLINE=91099948; PubMed=1987023;			
RA	Janson H., Heden L.-O., Grubb A., Ruan M., Forsgren A.;			
RT	"Protein D, an immunoglobulin D-binding protein of Haemophilus			
RT	influenzae: cloning, nucleotide sequence, and expression in			
RT	Escherichia coli";			
RL	Infect. Immun. 59:119-125(1991).			
RL	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Minna / Serotype B;			
RC	MEDLINE=94011360; PubMed=8104899;			
RA	Janson H., Ruan M., Forsgren A.;			
RT	"Limited diversity of the protein D gene (hpd) among encapsulated and			
RT	nonencapsulated Haemophilus influenzae strains.";			
RL	Infect. Immun. 61:4546-4552(1993).			
RL	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=3639, 3640, 6-7626, Eagan / Serotype B, HK695 / Serotype B, and			
RC	NCTC 8468 / Serotype B;			
RC	MEDLINE=95122210; PubMed=7822043;			

Db 299 MVQDAQONKLVVHPYTVRSKDLPEYTTVDVQVLYDALYNKAGVNGLFTDPDPKAVKFL 355

RESULT 7

Q8CV6 PRELIMINARY; PRT; 358 AA.

AC Q8CV6

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Glycerophosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46)

GN Name=glpQ; OrderedLocusNames=c2780;

OS Escherichia coli O6.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=217992;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

RA Mayhew G.P., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

DR EMBL: AE016763; AAN81234.1; -

DR GO: GO:0008889; F:glycerophosphodiester phosphodiesterase act. . . ; IEA.

DR GO: GO:0016787; F:hydrolase activity; IEA.

DR GO: GO:0006071; P:glycerol metabolism; IEA.

DR InterPro: IPR004129; GSDP.

DR Pfam: PF03009; GSDP; 1.

KW Complete proteome; Hydrolase.

SQ SEQUENCE 358 AA; 40826 MW; C633D35E8A86591F8 CRC64;

Query Match 50.6%; Score 1254.5; DB 2; Length 358;

Best Local Similarity 64.4%; Pred. No. 5.6e-79;

Matches 230; Conservative 55; Mismatches 65; Indels 7; Gaps 2;

QY 4 KTLALSLLAAGVLAACSSSHSNMANTQMSKIIIAHRGASGLPEHTLESKALAPAOQA 63

Db 6 KNLMSAIMMSTIVMGSSAAAD-----SNEKIVIAHRGASGLPEHTLPKAMAYAOQA 59

QY 64 DYLEDLAWTKDGLVWVTHDPLDGLTDVAKKFPHRHRKGRYVYIDFTLKEIQSLEWTE 123

Db 60 DYLEDQLVMTKDDHLVLDHLDVDRVDRPDRARCKGRYVYIDFTLKEIQSLEWTE 119

QY 124 NPTKDGKQAOVYVNRFLPWKSHFRIFHTFDEIEPIQGLEKSTGKKGVIYPIEIKAPWFFH 183

Db 120 GFDIENGKKVQYVPCGRFPMGKSDFRVHTFEIEIEFVQGLNHSHTGKNGIYPIEIKAPWFFH 179

QY 184 QNGKDIAETLTKVKKYGYDKTDMVYIQTDFNBLKRIKTELLPQMGMDLKVLIAYT 243

Db 180 QEGKDIAETLTKVKKYGYDKTDMVYIQTDFNBLKRIKTELLPQMGMDLKVLIAYT 239

QY 244 DWKETQEKDPKGYVYVNYNWMFKPGMAEVVKYADGPGWYMLVNEKSPDNIVVTP 303

Db 240 DWNETQQKQPGSWYVNYNWMFKPGMAEVVKYADGPGWYMLVNEKSPDNIVVTP 298

QY 304 LVKELAQVNEVHPYTVRSKDLPEYTTVDVQVLYDALYNKAGVNGLFTDPDPKAVKFL 360

Db 299 MVQDAQONKLVVHPYTVRSKDLPEYTTVDVQVLYDALYNKAGVNGLFTDPDPKAVKFL 355

RESULT 8

GLPQ ECOLI STANDARD; PRT; 358 AA.

AC P09394;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)

DE Glycerophosphoryl diester phosphodiesterase, periplasmic precursor (EC 3.1.4.46) (Glycerophosphodiester phosphodiesterase).

GN Name=glpQ; OrderedLocusNames=b2239;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-31.

RX MEDLINE=91238712; PubMed=1851953;

RA Tomassen J., Eiglmeier K., Cole S.T., Overduin P., Larson T.J.,

RA Boos W.;

RT "Characterization of two genes, glpQ and ugpQ, encoding glycerophosphoryl diester phosphodiesterases of Escherichia coli.";

RL Mol. Gen. Genet. 226:321-327(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97349980; PubMed=9205837;

RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,

RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,

RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,

RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,

RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,

RA Yanagata S., Horiuchi T.;

RT "Construction of a contiguous 874-kb sequence of the Escherichia coli-K12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features.";

RL DNA Res. 4:91-113(1997).

RN [4]

RP SEQUENCE OF 1-9 FROM N.A.

RC STRAIN=K12;

RX MEDLINE=88201663; PubMed=3329281;

RA Eiglmeier K., Boos W., Cole S.;

RT "Nucleotide sequence and transcriptional startpoint of the glpT gene of Escherichia coli: extensive sequence homology of the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system.";

RL Mol. Microbiol. 1:251-258(1987).

RN [5]

RP SEQUENCE OF 26-44.

RC STRAIN=K12;

RX PubMed=8899705;

RA Gonzalez-Gil G., Bringmann P., Kahmann R.;

RT "Fis is a regulator of metabolism in Escherichia coli.";

RL Mol. Microbiol. 22:21-29(1996).

CC -!- FUNCTION: Glycerophosphoryl diester phosphodiesterase hydrolyzes deacylated phospholipids to G3P and the corresponding alcohols.

CC -!- CATALYTIC ACTIVITY: A glycerophosphodiester + H(2)O = an alcohol + sn-glycerol 3-phosphate.

CC -!- SUBCELLULAR LOCATION: Periplasmic.

CC -!- MISCELLANEOUS: There are 2 isozymes of glycerophosphoryl diester phosphodiesterase in E.coli: a periplasmic isozyme (glpQ) and a cytosolic isozyme (ugpQ).

CC -!- SIMILARITY: Belongs to the glycerophosphoryl diester phosphodiesterase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@iab-sib.ch).

```
CC -----
CC EMBL; X56907; CAA40323.1; -.
DR EMBL; U00096; AAC75299.1; -.
DR EMBL; D90855; BAA16058.1; -.
DR EMBL; Y00536; CAA68599.1; -.
DR PIR; S15945; S15945.
DR ECODBASE; C039.0; 6TH EDITION.
DR ECODBASE; EB0394; -.
DR Ecogene; EG10399; gipQ.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome; Direct protein sequencing; Glycerol metabolism;
KW Hydrolase; Multigene family; Periplasmic; Signal.
FT SIGNAL 1 25
FT CHAIN 26 358
FT   Glycerophosphoryl diester
FT   phosphodiesterase, periplasmic.
FT   D -> I (in Ref. 5).
FT CONFLICT 27
SQ SEQUENCE 358 AA; 40843 MW; B08FD27399641616 CRC64;

Query Match          50.5%; Score 1251.5; DB 1; Length 358;
Best Local Similarity 64.1%; Pred. No. 9.1e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLGACSSHSNMANTQMSDKIIIAHAGSGYLPHTLESKALAFQAQA 63
Db 6 KNLSSAIMSTIVMGSSAAMD-----SNEKIVIAHAGSGYLPHTLPKAMAYAQGA 59

Qy 64 DYLEQDLAMTKDGRVLVVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEWTE 123
Db 60 DYLEQDLVMTKDDNLVVLHDLVDRVTDVADRFPPDRARKDGRYYAIDFTLDEIKSLKFT 119

Qy 124 NFETKDGKQAOQVYVNPFRPLMKSHPRHTFEDEIFIOGLEKSTGKKGVIYPIKAPWPHH 183
Db 120 GFDIENGGKVQYTPGRFPMGKSDPRVHTFEEIEFVQGLNHSSTGKNIYPIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLLKKGVDKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYT 243
Db 180 QEGKDIAAKTLEVLKKGYYTGKDDKVVYLCQFDADDELAKRIKNELEPKMGMLNLVQLIAYT 239

Qy 244 DWKETQEKDPKGYWVNYNDWMFKPGMAEVVYKADGVGPGWYMLVKNKESKPNIVVTP 303
Db 240 DWNETQKQPDGWSWVNYNDWMFKPGMAKQVAYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNQMDALLNKSGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLIVVHPYTVRSKPLPEYTPDVNQYLYDALYNKAGVNGLFTDFPDKAVKFL 355

RESULT 9
Q7UC72 PRELIMINARY; PRT; 358 AA.
AC Q7UC72;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycerophosphodiester phosphodiesterase.
GN Name=gipQ; OrderedLocNames=S2454;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RT
```

```
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE016986; AAP17654.1; -.
DR GO; GO:0008899; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
SQ SEQUENCE 358 AA; 40893 MW; ABCD1B395FE41D41 CRC64;

Query Match          50.5%; Score 1251.5; DB 2; Length 358;
Best Local Similarity 64.1%; Pred. No. 9.1e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLGACSSHSNMANTQMSDKIIIAHAGSGYLPHTLESKALAFQAQA 63
Db 6 KNLSSAIMSTIVMGSSAAMD-----SNEKIVIAHAGSGYLPHTLPKAMAYAQGA 59

Qy 64 DYLEQDLAMTKDGRVLVVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEWTE 123
Db 60 DYLEQDLVMTKDDHLVVLHDLVDRVTDVADRFPPDRARKDGRYYAIDFTLDEIKSLKFT 119

Qy 124 NFETKDGKQAOQVYVNPFRPLMKSHPRHTFEDEIFIOGLEKSTGKKGVIYPIKAPWPHH 183
Db 120 GFDIENGGKVQYTPGRFPMGKSDPRVHTFEEIEFVQGLNHSSTGKNIYPIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLLKKGVDKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYT 243
Db 180 QEGKDIAAKTLEVLKKGYYTGKDDKVVYLCQFDADDELAKRIKNELEPKMGMLNLVQLIAYT 239

Qy 244 DWKETQEKDPKGYWVNYNDWMFKPGMAEVVYKADGVGPGWYMLVKNKESKPNIVVTP 303
Db 240 DWNETQKQPDGWSWVNYNDWMFKPGMAKQVAYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNQMDALLNKSGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLIVVHPYTVRSKPLPEYTPDVNQYLYDALYNKAGVNGLFTDFPDKAVKFL 355

RESULT 10
Q83QU2 PRELIMINARY; PRT; 377 AA.
AC Q83QU2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycerophosphodiester phosphodiesterase, periplasmic.
GN Name=gipQ; OrderedLocNames=SF321;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang P., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
DR EMBL; AE015248; AAM43836.1; -.
DR GO; GO:0008899; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome.
SQ SEQUENCE 377 AA; 42920 MW; B4EA7855FACAA936 CRC64;

Query Match          50.5%; Score 1251.5; DB 2; Length 377;
Best Local Similarity 64.1%; Pred. No. 9.7e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;
```

```

QY 4 KTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGLYVPEHTLESKALAPAOQA 63
Db 25 KULSMAIMSTIVMGSSMAAD-----SNEKIVIAHRGASGLYVPEHTLPKAMAWAQA 78
QY 64 DYLEDQDLAMTKDGLRVLVHDFELDGLTDVAKKFPHRHRKDGRRYVVDFTLKEIQSLEWTE 123
Db 79 DYLEDQDLVMTKDDHLVLDHLDLDRVTDVADRPDRARCKGRYVVDFTLDEIKSLKFT 138
QY 124 NFETKDGKQAOQVYNNRFLPKWKSHPRIHTFEDEIEFIQGLEKSTGKVGIVPEIKAPWFFH 183
Db 139 GFDIENGKKVQVYPCRFPMGKSDFRVHTFEDEIEFVQGLNHSFGKNGIYPIEIKAPWFFH 198
QY 184 QNGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 199 QEGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 258
QY 244 DKNETQEKDPKGVWYNNYNDMMFKPGMAEVVKYAGVCGPMYMLVKNKEESKPDNIYVTP 303
Db 259 DNETQEKDPKGVWYNNYNDMMFKPGMAEVVKYAGVCGPMYMLVKNKEESKPDNIYVTP 317
QY 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDPTGVFL 360
Db 318 MVQDAQONKLVVHPYTVRSKLPETVDPVQVLYDLYNKAGVNGLFTDFPDKAVKFL 374
RESULT 11
Q6CZ14
ID Q6CZ14 PRELIMINARY; PRT; 361 AA.
AC Q6CZ14;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Glycero-phosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46).
GN Name=glpQ; OrderedLocustNames=ECA4167;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmon G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX50851; CAG77064.1; -.
DR GO; GO:0008889; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GSDP.
DR Pfam; PF03009; GSDP; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 361 AA; 41551 MW; EB3E20B4555AB5F6 CRC64;
Query Match 50.4%; Score 1248; DB 2; Length 361;
Best Local Similarity 63.9%; Pred. No. 1.6e-78;
Matches 228; Conservative 55; Mismatches 70; Indels 4; Gaps 2;
QY 5 TSLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGLYVPEHTLESKALAPAOQA 64
Db 6 TSLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGLYVPEHTLPKAMAWAQA 62
QY 65 YLEQDLAMTKDGLRVLVHDFELDGLTDVAKKFPHRHRKDGRRYVVDFTLKEIQSLEWTE 124

```

```

Db 63 YLEQDLAMTKDGLRVLVHDFELDGLTDVAKKFPHRHRKDGRRYVVDFTLKEIKSLKFTSG 122
QY 125 FETKDGKQAOQVYNNRFLPKWKSHPRIHTFEDEIEFIQGLEKSTGKVGIVPEIKAPWFFH 184
Db 123 PDKNGKQVQSVYNNRFLPKWKSHPRIHTFEDEIEFIQGLNHSFGKNGIYPIEIKAPWFFH 182
QY 185 NGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 244
Db 183 EGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 242
QY 245 WKETQEKDPKGVWYNNYNDMMFKPGMAEVVKYAGVCGPMYMLVKNKEESKPDNIYVTP 304
Db 243 WHETQEKADGSWYNNYNDMMFKPGMAEVVKYAGVCGPMYMLVKNKEESKPDNIYVTP 301
QY 305 VKELAQYNVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDPTGVFL 361
Db 302 VKEAHQHLKVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDPTGVFL 358
RESULT 12
Q8ZNG7
ID Q8ZNG7 PRELIMINARY; PRT; 356 AA.
AC Q8ZNG7;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Periplasmic glycerophosphodiester phosphodiesterase (EC 3.1.4.46).
GN Name=glpQ; OrderedLocustNames=STM2282;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008802; AAL21183.1; -.
DR GO; GO:0008889; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR Pfam; PF03009; GSDP; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 356 AA; 40425 MW; 527261A667E09C44 CRC64;
Query Match 49.9%; Score 1235; DB 2; Length 356;
Best Local Similarity 64.5%; Pred. No. 1.3e-77;
Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;
QY 4 KTLALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGLYVPEHTLESKALAPAOQA 63
Db 6 KNLVSVALMAGMTGSGA-----VAAEKVIAHRGASGLYVPEHTLPKAMAWAQA 56
QY 64 DYLEDQDLAMTKDGLRVLVHDFELDGLTDVAKKFPHRHRKDGRRYVVDFTLKEIQSLEWTE 123
Db 57 DYLEDQDLVMTKDDHLVLDHLDLDRVTDVADRPDRARCKGRYVVDFTLDEIKSLKFT 116
QY 124 NFETKDGKQAOQVYNNRFLPKWKSHPRIHTFEDEIEFIQGLEKSTGKVGIVPEIKAPWFFH 183
Db 117 GFDIENGKKVQVYPCRFPMGKSDFRVHTFEDEIEFVQGLNHSFGKNGIYPIEIKAPWFFH 176
QY 184 QNGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 177 QEGKDIAETLKVLYKGYDKTDMVYLOTDFPNELKRIKTELLPQMGMDLKLVLQIAYT 236

```


Search completed: June 22, 2005, 11:40:43
Job time : 175 secs

THIS PAGE BLANK (USPTO)